WILEY-VCH

Sensors Update

Edited by
H. Baltes, G. K. Fedder, J. G. Korvink

Volume 12

Sensors Technology – Applications – Markets

MICHIGAN MOLECULAR INSTITUTE
31227346

Dear Reader,

Sensors Update is also available **online** at Wiley InterScience!

Interested? Please check out *http://www.interscience.wiley.com*

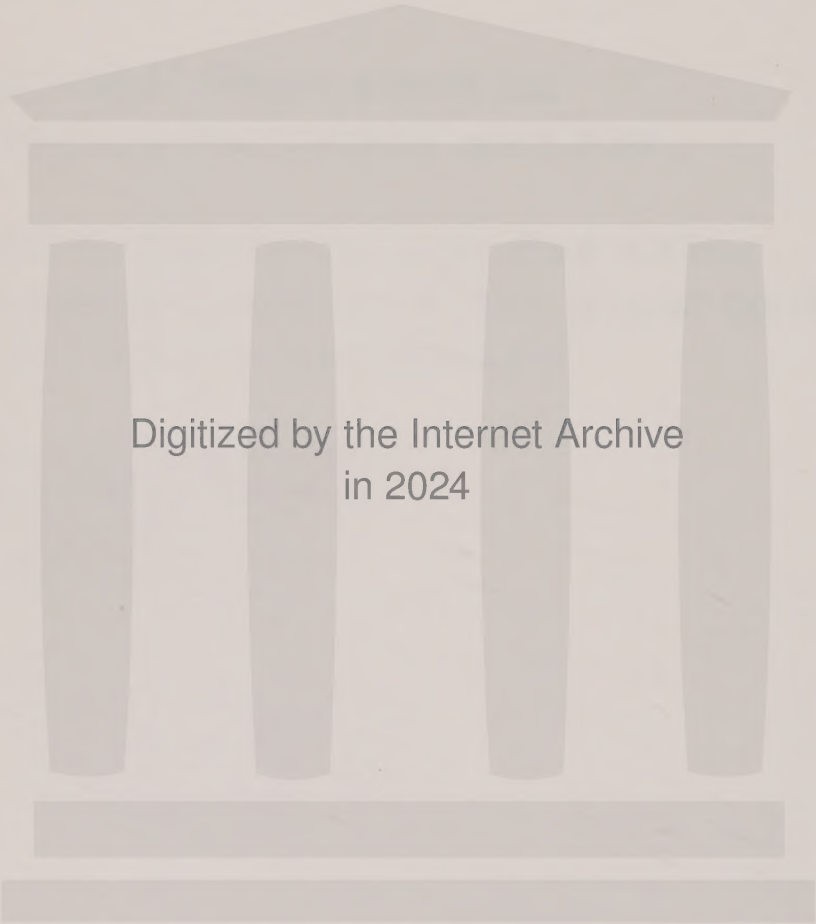
For information how to get **free online access** at no additional cost, please contact us according to your location!

Germany, Switzerland, Austria, Luxembourg, Liechtenstein:
service@wiley-vch.de
Tel. +49 (0)6201 606 400
Fax +49 (0)6201 606 184

All other countries:
E-mail: *cs-journals@wiley.co.uk*
Tel. (UK only) 0800 243407
(overseas) +44 1243 843294
Fax +44 1243 843296

MICHIGAN MOLECULAR INSTITUTE
1910 WEST ST. ANDREWS ROAD
MIDLAND, MICHIGAN

MICHIGAN MOLECULAR INSTITUTE
1910 WEST ST. ANDREWS ROAD
MIDLAND, MICHIGAN 48640


Digitized by the Internet Archive
in 2024

https://archive.org/details/sensorsupdate12s0000unse

Sensors Update

Volume 12

Sensor Technology –
Applications – Markets

Related Wiley-VCH titles:

Sensors Vol. 1–9
edited by W. Göpel, J. Hesse, J.N. Zemel

ISBN 3-527-26538-4

Sensors Applications
edited by J. Hesse, J.W. Gardner, W. Göpel

Vol. 1

Sensors in Manufacturing
edited by H.K. Tönshoff, I. Inasaki

ISBN 3-527-29558-5

Vol. 2

Sensors in Intelligent Buildings
edited by O. Gassmann, H. Meixner

ISBN 3-527-29557-7

Upcoming Volumes:

Sensors in Automotive Technology
Sensors in Household Appliances
Sensors in Medicine and Health Care
Sensors in Aerospace Technology
Sensors in Environmental Technology

Sensors Update

Edited by
H. Baltes, G. K. Fedder, J. G. Korvink

Volume 12

MICHIGAN MOLECULAR INSTITUTE
1910 WEST ST. ANDREWS ROAD
MIDLAND, MICHIGAN 48640

WILEY-VCH GmbH & Co. KGaA

Series Editors

Prof. Dr. H. Baltes
Physical Electronics Laboratory
ETH Hoenggerberg, HPT-H6
8093 Zürich
Switzerland

Prof. Gary K. Fedder
ECE Department &
The Robotics Institute
Carnegie Mellon University
Pittsburgh, PA 15213-3890
USA

Prof. Dr. Jan G. Korvink
Institute for Microsystem
Technology
University of Freiburg
Georges-Köhler-Allee 103
79110 Freiburg
Germany

The series Sensors Update was founded by
Prof. Dr. H. Baltes, Prof. Dr. W. Göpel and Prof. Dr. J. Hesse.

This book was carefully produced. Nevertheless, authors, editors, and publisher do not warrant the information contained therein to be free of errors. Readers are advised to keep in mind that statements, data, illustrations, procedural details or other items may inadvertently be inaccurate.

Library of Congress Card No.: applied for

British Library Cataloguing-in-Publication Data:
A catalogue record for this book is available from the British Library

Bibliographic information published by Die Deutsche Bibliothek
Die Deutsche Bibliothek lists this publication in the Deutsche Nationalbibliografie;
detailed bibliographic data is available in the Internet at <http://dnb.ddb.de>.

ISBN 3-527-30602-1

© 2003 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

Printed on acid-free paper.

All rights reserved (including those of translation in other languages). No part of this book may be reproduced in any form – by photoprinting, microfilm, or any other means – nor transmitted or translated into machine language without written permission from the publishers. Registered names, trademarks, etc. used in this book, even when not specifically marked as such, are not to be considered unprotected by law.

Composition: K+V Fotosatz GmbH, Beerfelden
Printing: betz-druck GmbH, Darmstadt
Bookbinding: J. Schäffer GmbH & Co. KG, Grünstadt
Printed in the Federal Republic of Germany

793
6

Preface

Sensors Update is a book series that provides timely reviews in the sensors and microsystems area. This 12th volume of *Sensors Update* series follows the successful nine-volume series *Sensors*. The three major sections of each volume are 'Sensor Technology', where key areas in applied and basic research are reviewed, 'Sensor Applications', which is a forum for up-to-date overviews of application areas, and 'Sensor Markets', where emerging trends, reviews of suppliers, or patent status for a particular market sector are reviewed. *Sensors Update* covers topics spanning fundamental concepts, theory and simulation, materials, devices, system applications and commercial market aspects.

This 12th volume begins with technology articles on environmental sensors, integrated gas detectors, and X-ray imaging. The next two articles detail recent advances in broad and important application areas, micro-optical systems and RF-microelectromechanical systems. This volume is rounded out with two sensor market overviews in the automotive and the household appliance sectors.

This current volume of *Sensors Update* has been produced due to the timely effort of many people. First and foremost, the series editors thank the authors for a wonderful set of overview papers. The editors are grateful to the publishers, Wiley-VCH, for their long-term support of the series. In particular, the editors thank the series publishing editor, Dr. Martin Ottmar, for his management of this enterprise and the staff Hans-Jochen Schmitt and Dr. Jörn Ritterbusch for their support in editing the articles.

H. Baltes, Zürich G.K. Fedder, Pittsburgh J.G. Korvink, Freiburg

Editorial Advisory Board

M. Allen, Georgia Tech, Atlanta, USA
R.M. Crooks, Texas A&M University, College Station, USA
M. Esashi, Tohoku University, Sendai, Japan
H. Fujita, University of Tokyo, Tokyo, Japan
J. Gardner, Warwick University, Coventry, UK
G. Guilbault, University College Cork, Cork, Ireland
G.T.A. Kovacs, Stanford University, Stanford, USA
I. Lundström, Linköpings Universitet, Linköping, Sweden
H. Meixner, Siemens AG, Munich, Germany
N. Miura, Kyushu University, Fukuoka, Japan
T. Moriizumi, Tokyo Institute of Technology, Tokyo, Japan
O. Paul, Institute for Microsystem Technology, Freiburg, Germany
D.N. Reinhoudt, Twente University, Enschede, The Netherlands
A.J. Ricco, ACLARA BioSciences, Mountain View, USA
N.F. de Rooij, Institute of Microtechnology, Neuchâtel, Switzerland
J.R. Stetter, Illinois Institute of Technology, Chicago, USA
U. Weimar, University of Tübingen, Tübingen, Germany
H. Zappe, Institute for Microsystem Technology, Freiburg, Germany

Contents

List of Contributors IX

Part 1 Sensor Technology

1.1 Large Area Digital X-ray Imaging 3
S. Tao, K.S. Karim, P. Servati, C.-H. Lee, A. Nathan

1.2 CMOS Single-chip Gas Detection Systems: Part II 51
C. Hagleitner, A. Hierlemann, H. Baltes

1.3 Ionomer-coated Electrodes and Nanoelectrode Ensembles as Electrochemical Environmental Sensors: Recent Advances and Prospects 121
P. Ugo, L.M. Moretto, F. Vezzà

Part 2 Sensor Applications

2.1 Modular Microoptical Systems for Sensors and Telecommunication . . 143
U. Wallrabe, J. Mohr

2.2 MEMS in RF Filter Applications: Thin-film Bulk Acoustic Wave Technology 175
R. Aigner

Part 3 Sensor Markets

3.1 Automotive Sensor Market 213
K.-F. Reinhart, M. Illing

3.2 Market Data – Sensors in Household Appliances 231
G.R. Tschulena

Index 243

List of Contributors

Robert Aigner
Infineon Technologies AG
Secure Mobile Solutions
TI TM MI
Otto-Hahn-Ring 6
81739 Munich
Germany

Henry Baltes
Physical Electronics Lab
ETH Hoenggerberg, HPT-H6
CH-8093 Zurich
Switzerland

Christoph Hagleitner
Physical Electronics Lab
ETH Hoenggerberg, HPT-H3
CH-8093 Zurich
Switzerland

Andreas Hierlemann
Physical Electronics Lab
ETH Hoenggerberg, HPT-H4.2, IQE
CH-8093 Zurich
Switzerland

Matthias Illing
Robert Bosch GmbH
AE/TPS-I
72703 Reutlingen
Germany

Karim S. Karim
Electrical and Computer Engineering
University of Waterloo
Waterloo, Ontario N2L 3G1
Canada

Czang-Ho Lee
Electrical and Computer Engineering
University of Waterloo
Waterloo, Ontario N2L 3G1
Canada

Jürgen Mohr
Forschungszentrum Karlsruhe
Institut für Mikrostrukturtechnik
Postfach 3640
76021 Karlsruhe
Germany

Ligia Maria Moretto
Department of Physical Chemistry
University of Venice
Calle Larga S.Marta 2137
I-30123 Venice
Italy

Arokia Nathan
Electrical and Computer Engineering
University of Waterloo
Waterloo, Ontario N2L 3G1
Canada

Karl-Franz Reinhart
Robert Bosch GmbH
FV/FLD
70839 Gerlingen
Germany

Peyman Servati
Electrical and Computer Engineering
University of Waterloo
Waterloo, Ontario N2L 3G1
Canada

Sheng Tao
Electrical and Computer Engineering
University of Waterloo
Waterloo, Ontario N2L 3G1
Canada

Guido Tschulena
sgt Sensor Consulting
Reichenberger Str. 5
61273 Wehrheim
Germany

Paolo Ugo
Department of Physical Chemistry
University of Venice
Calle Larga S.Marta 2137
I-30123 Venice
Italy

Francesca Vezzà
Department of Physical Chemistry
University of Venice
Calle Larga S.Marta 2137
I-30123 Venice
Italy

Ulrike Wallrabe
Forschungszentrum Karlsruhe
Institut für Mikrostrukturtechnik
Postfach 3640
76021 Karlsruhe
Germany

PART 1

Sensor Technology

1.1 Large Area Digital X-ray Imaging

S. Tao, K. S. Karim, P. Servati, C.-H. Lee and A. Nathan,
University of Waterloo, Waterloo, ON, Canada

Abstract

This chapter reviews amorphous silicon devices for large area flat panel imaging technology. We present Schottky and p-i-n diode image sensors and elaborate on their operating principles, electrical and optoelectronic characteristics including stability, along with the challenges associated with reduction of the dark current. Issues pertinent to sensor-thin film transistor integration for different active matrix pixel architectures for high fill factor imaging arrays are presented along with optimization of materials and processing conditions for reduced threshold voltage shift, reduced parasitics and leakage current, and enhanced mechanical integrity. Extension of the current fabrication processes to low ($\sim 120\,^{\circ}C$) temperature, enabling fabrication of flexible imaging array (on plastic substrates), is also discussed.

Keywords: Schottky image sensors; p-i-n diode image sensors; active pixel sensors; a-Si technology

Contents

1.1.1 Introduction 4
1.1.1.1 Active Matrix Arrays 5
1.1.1.2 Conventional Imaging Methods 5
1.1.1.3 Detection Schemes for Flat Panel X-ray Imaging 6
1.1.2 Material Considerations 7
1.1.2.1 Amorphous Silicon 8
1.1.2.2 Dielectric Materials 10
1.1.2.3 Metals 12
1.1.3 Detectors 15
1.1.3.1 Photoconductors 15

1.1.3.2 Mo/a-Si:H and ITO/a-Si:H Schottky Diodes 16
1.1.3.3 p-i-n Photodiodes 23

1.1.4 Pixel Architectures and Integration 27
1.1.4.1 Thin Film Transistor and Leakage Current 27
1.1.4.2 Active Pixel Sensors 29
1.1.4.3 Pixel Integration 32
1.1.4.4 Phosphor Integration 35

1.1.5 Imaging Arrays 39
1.1.5.1 Readout Operation 40
1.1.5.2 Capacitance Extraction for Large Area Arrays 41

1.1.6 New Challenges in Large Area Digital Imaging 44

1.1.7 References 45

1.1.1 Introduction

Pixelated arrays of electronics in amorphous silicon (a-Si:H) technology, routinely used for liquid crystal displays, are now being extended to several new and significant application areas in large area digital imaging [1, 2]. Large area active matrix flat panel imagers (AMFPIs) with area $\sim 30 \times 40\ cm^2$ have been demonstrated for radiography [3]. Interest in a-Si:H technology stems from a variety of desired materials and technological attributes [4]. The high optical absorption, low temperature deposition ($<300\,^{\circ}C$), high uniformity over a large area, few constraints on substrate size, material, or topology, standard integrated circuit (IC) lithography processes, and low capital equipment cost, associated with the a-Si:H material, offer a viable technological alternative for improved imaging of optical signals and high energy radiation. Notable application areas include contact imaging for document scanning, digital copiers, and fax machines; color sensors/imaging; position/motion detection; radiation detection/imaging of x-rays in biomedical applications, gamma-ray space telescopes, airport security systems, and non-destructive testing of mechanical integrity of materials or structures.

Extensive reviews on applications of the a-Si:H flat panel imager technology to biomedical x-ray imaging can be found in a recently published *Handbook of Medical Imaging* [5], which describes the physics of the different imaging modalities. We will focus on Schottky and p-i-n diode image sensors and describe their operating principles, materials-related and processing issues, electrical and optoelectronic characteristics and stability, along with the new challenges that lie ahead for reduction of dark current. Also, issues pertinent to sensor–thin film transistor (TFT) integration will be discussed along with new pixel architectures for high fill factor imaging arrays with reduced parasitic capacitance, together with processing conditions for reduced threshold voltage (V_T) shift and leakage current, and enhanced mechanical integrity. Selected results are shown for x-ray

and optical detectors and integrated x-ray pixel structures. Extension of the current fabrication processes to low temperature ($\sim 120\,^{\circ}\mathrm{C}$), permitting the fabrication of flexible (curved) imagers (on polymeric substrates) for high light collection efficiency, is also discussed along with preliminary results in terms of the static characteristics of the active matrix switch.

1.1.1.1 Active Matrix Arrays

There are two architectures currently employed in large area AMFPIs: the linear architecture, which is used in photocopiers, fax machines, and scanners, and the two-dimensional or area array architecture, which is employed in digital (including video) lensless cameras and x-ray imaging systems. In both linear and 2-D architectures, the basic imaging unit is the pixel, which consists of an image sensor and a switch. The pixel is accessed by a matrix of gate and data lines, and operated in storage (or integration) mode. Here, during the off-period of the switch, the sensor charge is integrated, and when the switch is turned on, the charge in the sensor is transferred to the data line where it is detected by a charge-sensitive amplifier. Currently either the diode [6] or the TFT [7] is used for the switching element. Although the diode has fewer masking steps and simpler connectivity, it comes with high capacitance and non-linear current–voltage characteristics. In contrast, the TFT has a lower capacitance, linear switching characteristics, and low leakage current, and is thus more widely used despite its potentially large shift in V_{T} after prolonged gate bias. The reduction of both leakage and instability remains a key challenge from the standpoint of material/processing and device design [8].

1.1.1.2 Conventional Imaging Methods

Wilhelm Roentgen created the first x-ray image (of his wife's hand) in 1895. Since then, thousands of physicians and radiologists have developed various methods to acquire and interpret x-ray images. The most widely used method to detect x-rays is to convert the radiation to visible light through a phosphor screen for subsequent detection by a light-sensitive film [5]. Modern screen–film systems typically employ one or a pair of phosphor screens in combination with the light-sensitive film packaged in a light-tight cassette. This system, which has been accepted by radiologists for nearly a century, provides good quality x-ray images. The main drawbacks of the screen–film system are storage costs and digital incompatibility. Another method that is widely used, particularly in fluoroscopic imaging, is the image-intensifier tube, which is integrated with a charge-coupled device (CCD) or complementary metal oxide semiconductor (CMOS) camera. Here, x-rays first interact with the input phosphor to produce optical photons, which in turn eject electrons from the photocathode. The electrons are

accelerated across the vacuum tube and are focused on an output phosphor to produce optical photons, which are then viewed by a camera. This method employs multiple transduction steps to convert x-rays to visible light, which compromises image quality. Moreover, the complex lens and mechanical systems and the assembled arrays of CCD or CMOS lead to cost and size issues. As an alternative to these conventional detection methods, large area active matrix flat panel digital x-ray imagers based on a-Si:H technology have been developed over the past decade [9]. The motivation for AMFPIs include improved image quality, large area imaging capability, flat panel structure, low storage costs, and computerized handling/storage of sensory information.

1.1.1.3 Detection Schemes for Flat Panel X-ray Imaging

The detection schemes for x-rays can generally be divided into two categories: the direct detection scheme, where the x-rays are directly absorbed and converted to electrical charge in the detector; and the indirect detection scheme, where the x-rays are first converted into visible light by a phosphor layer, which in turn is converted to electrical charge in the photodetector. The electrical charge is read out by means of an a-Si:H active matrix (TFT) array.

In the direct detection scheme, a thick photoconductor, such as amorphous selenium (a-Se), is deposited over the a-Si:H active matrix [10], as depicted in Figure 1.1.1a. Incident x-rays are absorbed in the a-Se film and directly converted to electron–hole (e–h) pairs, which are then collected by an electric field

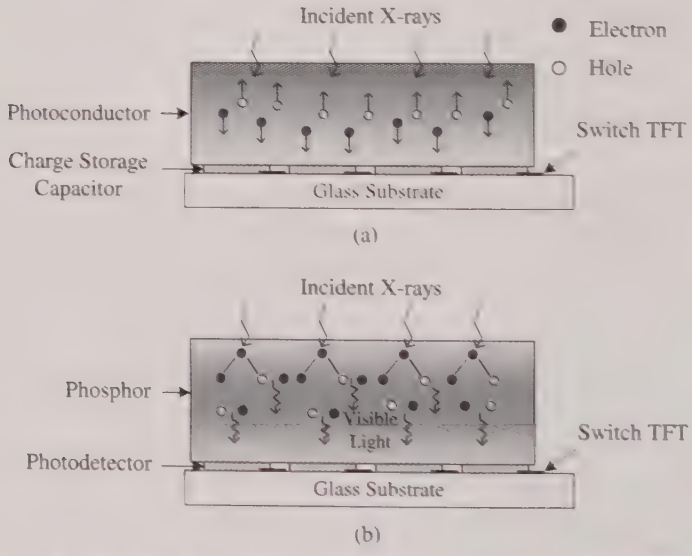


Figure 1.1.1. (**a**) Direct and (**b**) indirect detection schemes for large area digital x-ray imaging.

Table 1.1.1. Requirements for digital diagnostic medical imaging arrays (adapted from [5])

	Radiography	Mammography	Fluoroscopy
Imager size (cm)	35×43	18×24	25×25
Pixel area (μm^2)	200×200	50×50	250×250
Pixel count	1750×2150	3600×4800	1000×1000
Image readout time (s)	<5	<5	0.033/frame
X-ray spectrum (kV_p)	120	30	80
Exposure range (mR)	0.03–3	0.6–240	0.0001–0.01

that is applied across the photoconductor layer. A high electric field (~ 10 V/μm) is required to separate the e–h pairs. Normally, a 500 μm thick a-Se layer is required to absorb all the x-rays, and a bias voltage as high as 5–10 kV is applied across the a-Se film for efficient charge collection. Another approach uses a heavy metal Mo/a-Si:H Schottky diode as the x-ray detector [11], although its detective quantum efficiency is comparably much lower. The incident x-rays are absorbed and converted to high energy electrons inside the Mo layer by virtue of the photoelectric effect. These electrons are injected into the a-Si:H layer, and e–h pairs are then generated in the a-Si:H by avalanche multiplication. This detection scheme is operated at a low bias (<5 V).

The indirect detection scheme, as shown in Figure 1.1.1b, relies on a phosphor layer that is either assembled or integrated with the photodetector. The phosphor film absorbs the incident x-ray photons and produces high energy electrons, which in turn generate many e–h pairs in the phosphor. Visible light is emitted from the phosphor when the electrons and holes recombine, and detected by a-Si:H image sensor arrays located beneath the phosphor film.

Irrespective of the detection scheme, the requirements on x-ray spectrum and imager exposure for the different modalities are determined from those currently used in clinical practice [5]. These design requirements are summarized in Table 1.1.1. Here, mammography and fluoroscopy are considered particularly challenging because of the small (50×50 μm^2) pixel area requirements for mammography and because of the extremely low exposure range (μR) in fluoroscopy.

1.1.2 Material Considerations

The fabrication of AMFPIs is based on a-Si:H thin film technology. Materials used in array fabrication include amorphous silicon and associated dielectrics, and metals, which are deposited using plasma enhanced chemical vapor deposition (PECVD) and physical vapor deposition (PVD), respectively. The PECVD technique provides uniform device quality amorphous silicon and dielectric films

over a large area with low thermal budget. In the PECVD process, the reactive gases are dissociated by the plasma and then adsorbed by the substrate where the films are grown. Process parameters such as temperature, power, pressure, and reactive gas flow, influence the film properties and thus have a direct impact on the device characteristics. The metal films are deposited using rf magnetron sputtering, which provides for high surface smoothness and film uniformity over a large area. In the sputtering process, the atoms of the target are ejected from the surface under the bombardment of high energy inert gas ions (such as Ar^+). The ejected atoms travel through a certain distance to the substrate and condense on the surface of the substrate. The various materials used in the array fabrication are discussed in the following.

1.1.2.1 Amorphous Silicon

The disorder in atomic structure distinguishes a-Si:H from crystalline silicon (c-Si) [12]. The material has a short range order, which leads to an electric structure similar to c-Si (see Figure 1.1.2) However, instead of abrupt band edges, the a-Si:H has broadened tails of states extending into the forbidden gap. These band tails originate from the long range bonding disorder, and affect the electronic transport in the material. Tunneling transitions occur in these localized states, which provide for hopping conduction in the band tail. Meanwhile, the coordination defects generate deep defect states inside the band gap. The deep defects change the electronic properties of the material by controlling the trapping and recombination of carriers. The incorporation of hydrogen in the amorphous silicon film ensures the passivation of some dangling bonds and reduces the density of deep defects. Commonly, the inclusion of 10% hydrogen in the film is considered as an optimum ratio to achieve high quality amorphous silicon films.

Owing to the disorder in atomic structure, the momentum of the optical band-to-band transitions is no longer conserved in a-Si:H. Therefore, a-Si:H possesses a direct optical band gap of approximately 1.7 eV. It has a larger optical

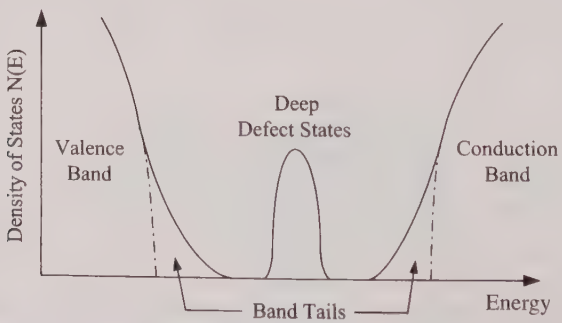


Figure 1.1.2. Schematic diagram of the density of states in a-Si:H. The dashed curves are the equivalent density of states in a crystal (adapted from [12]).

absorption coefficient in the visible range than c-Si, which makes it attractive for optical detection. Upon illumination, carriers are excited from the nonconducting to conducting states, and these photogenerated carriers in the band edges drift by virtue of the applied electric field towards the contacts. The recombination of carriers limits the photoconductivity of a-Si:H owing to charge trapping in the defect states. Therefore, the photoconductivity of a-Si:H is inversely proportional to the density of defect states.

In addition, external excitation such as light illumination and current flow can cause a reversible change in the density of states. This phenomenon is called 'metastability'. The most widely studied metastability in a-Si:H is the Staebler-Wronski effect, in which additional localized states are created by light illumination [13]. This effect decreases the photoconductivity and increases the dark conductivity of a-Si:H. These metastable defects can be removed by annealing the material at 150–200 °C.

The electrical and optical properties of a-Si:H are largely determined by the structural and coordination defects, dangling bonds and internal voids, and the effects of impurity atoms in the material. Impurities such as oxygen (O), nitrogen (N), and carbon (C) are inevitably incorporated into the film during deposition. The most dominant impurity incorporation mechanism stems from the wall interaction process [14, 15]. During deposition, reactive species in the plasma (eg, H) pick up impurity atoms from the wall in the gaseous phase. The gaseous impurities are dissociated in the plasma, and then re-incorporated into the film. This mechanism is of concern in single chamber PECVD systems owing to the cross-contamination of the various deposition processes. Besides the desorption of contaminants on the reactor wall, oil back-diffusion from the vacuum pump and impurities in the feed gases are also major sources of impurities. Typically, films made in an ultra-high vacuum (UHV) system typically contain 10^{17}–10^{18} cm^{-3} of O, 10^{17}–10^{18} cm^{-3} of C, and 10^{16}–10^{17} cm^{-3} of N. Because the atomic structure of a-Si:H is random, the incorporation of impurity atoms in the a-Si:H matrix is complex. It has been reported that the charged Si dangling bonds increase with the O or N incorporation inside the a-Si:H. There are two ways to explain the phenomenon. One is the so-called 'donor' model [16]. The N atom is in the form of N_4^+, which is positively charged fourfold-coordinated N. The O atom is in the form of O_3^+, which is positively charged threefold-coordinated O. In this case, the N or O contributes an electron to the a-Si:H. Owing to charge neutrality, a negatively charged Si dangling bond (DB) is generated. The other possible explanation follows from the 'potential fluctuations' model [17]. In this model, incorporated impurities generate structural randomness, and bring about fluctuations in the net-electron density at the Si sites, which in turn cause potential fluctuations. These fluctuations eventually result in an increase in the charged dangling bond density. In general, an increment of O, N, and C increases the defect density in the a-Si:H film, which in turn degrades the performance of devices such as photodiodes and TFTs.

Highly doped n^+a-Si:H is used to form an ohmic contact between the metal and a-Si:H layer. This can be achieved by adding a small quantity of phosphine to the process gas during deposition. The conductivity increases with doping lev-

el up to 1% phosphine to silane gas ratio, and then drops dramatically owing to the high density of defect states created at very high doping levels [12]. The doping mechanism can be explained in terms of substitutional doping. Because of the random amorphous network, the incorporation of dopants includes both threefold inactive dopants and fourfold substitutional dopants. The doping efficiency is defined as the ratio of active fourfold dopants to the total number of dopants. The n^+a-Si:H has a low doping efficiency and its resistivity is higher than 100 Ω-cm. To obtain a high doping efficiency and low resistivity, hydrogenated microcrystalline silicon (μc-Si:H) has been introduced. Hydrogen dilution plays an important role in the growth of μc-Si:H. The hydrogen radicals in the plasma act as an etchant of Si and promote chemical equilibrium between the deposition and etching of the growing surface. The growth process of μc-Si:H films is represented by a three layer model. During the initial nucleation stage, the film is mostly amorphous ($\sim$150 Å). Then microcrystallites start to form, in which the film is made of crystalline grains with amorphous grain boundaries. When the film grows thicker (>1000 Å), the growing surface becomes largely crystalline. Heavily phosphorus-doped hydrogenated microcrystalline silicon (n^+μc-Si:H) has a relatively high doping efficiency. Its resistivity can be as low as 0.1 Ω-cm [18].

1.1.2.2 Dielectric Materials

In imaging array applications, dielectric materials, such as a-SiN_x:H, a-SiO_x, and a-SiON, serve as gate insulators for TFTs and as masking materials for patterning. The a-SiN_x:H film is deposited by the dissociation of a gas mixture of NH_3, N_2, and SiH_4 in a PECVD chamber. The ideal structure of a-SiN_x:H is a 3:4 (Si:N)-coordinated random network. The composition of a-SiN_x:H is determined by the gas ratios of NH_3 and SiH_4 during the deposition, as shown in Table 1.1.2 [8]. Plasma deposited a-SiN_x:H is a wide band gap material with an energy of $\sim$5.3 eV.

Figure 1.1.3 illustrates schematically the density of states in a-SiN_x:H. The Si dangling bond generates gap states near the midgap, whereas the negative N^- center and Si–Si bond introduce gap states just above the valence band edge [19]. The density of defect states decreases when the a-SiN_x:H changes from Si-rich to N-rich, which can be explained by the reduction of Si dangling bonds in

Table 1.1.2. Changes in composition in a-SiN_x:H films for different NH_3/SiH_4 ratios

Gas ratio (NH_3/SiH_4)	x(N/Si)
5	1.1
10	1.39
20	1.5
25	1.52

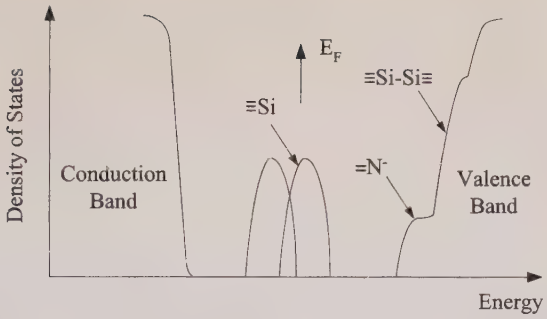


Figure 1.1.3. Schematic diagram of the density of states in plasma deposited a-SiN_x:H (≡Si, silicon dangling bond; $=N^-$, negative nitrogen dangling bond; ≡Si–Si≡, silicon–silicon weak bond) (adapted from [19]).

the N-rich a-SiN_x:H. The a-SiN_x:H film can be in either compressive or tensile stress depending on composition. It is in tensile stress when it is N-rich [8].

With ultra-thin a-SiN_x:H (<100 Å) materials needed for MIS photodiode structures, the initial growth of the a-SiN_x:H film plays an important role. At the initial stages of film deposition, nucleation happens at discrete spots on the surface, and then expands to the whole surface to form a continuous layer. The inhomogeneity of the initial nucleation causes the generation of pinholes and voids in the film, especially in an ultra-thin film of several atomic layers. Figure 1.1.4 illustrates transmission electron microscope (TEM) pictures of the tin-doped indium oxide (ITO)/a-SiN_x:H/a-Si:H interface at different a-SiN_x:H deposition times. With no a-SiN_x:H, a clear interface between the ITO and the a-Si:H is observed, as shown in Figure 1.1.4a. In the first 10 s of deposition, there is no obvious a-SiN_x:H layer between the ITO and the a-Si:H, as depicted in Figure 1.1.4b. However, roughness of the ITO/a-Si:H interface is observed, which indi-

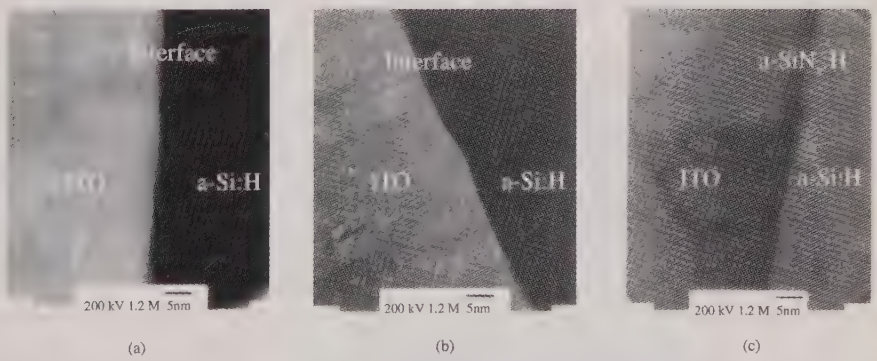


Figure 1.1.4. TEM pictures of the ITO/a-SiN_x:H/a-Si:H cross-section: (**a**) with no a-SiN_x:H; (**b**) after the first 10 s of a-SiN_x:H deposition; and (**c**) after 20 s of a-SiN_x:H deposition.

cates an inhomogeneous growth of a-SiN_x:H at the initial stages. After a 20 s deposition time, an a-SiN_x:H layer is observed between the ITO and a-Si:H. The thickness of the a-SiN_x:H layer is ~35 Å (see Figure 1.1.4c).

1.1.2.3 Metals

In imaging arrays, refractory metals, such as molybdenum (Mo), chromium (Cr), and nickel (Ni), are used as the gate metal in TFTs and as the back electrode metal in photodiodes. These metals have good thermal stability and high surface smoothness. Aluminum (Al), which has a high conductivity and good adhesion, serves as the top level interconnects and pad metallization. Owing to poor thermal stability, Al suffers from problems with hillock generation induced by subsequent high temperature process steps, as shown in Figure 1.1.5. For smooth Al, the process conditions, such as substrate temperature, process pressure, and rf power, have to be optimized.

In Mo/a-Si:H Schottky diode applications, the Mo layer (typically 500 nm thick) serves as the x-ray conversion layer. However, the internal stress in the Mo can be large, causing the film to peel off, especially in stacked structures. Figure 1.1.6 depicts the stress in Mo films deposited at different sputtering parameters on an intrinsic a-Si:H film [20]. The intrinsic stress of the Mo film can be either compressive or tensile, depending on the deposition pressure and the rf sputtering power. An optimum sputtering process is chosen to achieve low intrinsic stress.

In optical imaging applications, ITO film is often used as a transparent metal contact layer. It also serves as Schottky barriers in image sensors. ITO is a degenerate n-type wide-gap semiconductor with relatively low resistivity and high transparency in the visible range. The typical band gap of ITO is > 3.75 eV. The structure of ITO is based on the tin (Sn) substitution of indium (In) in the cubic bixbyite structure of indium oxide (In_2O_3). The Sn forms an interstitial bond with oxygen

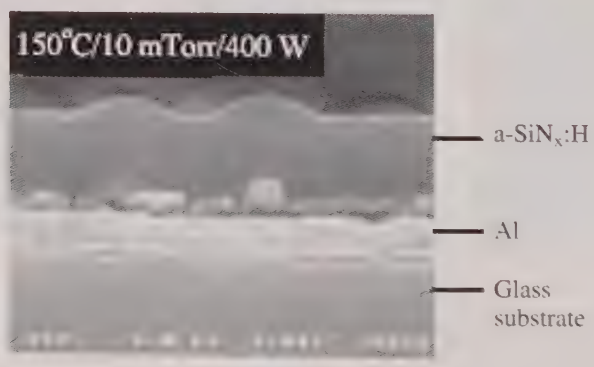


Figure 1.1.5. SEM cross section of Al/a-SiN interface with hillocks.

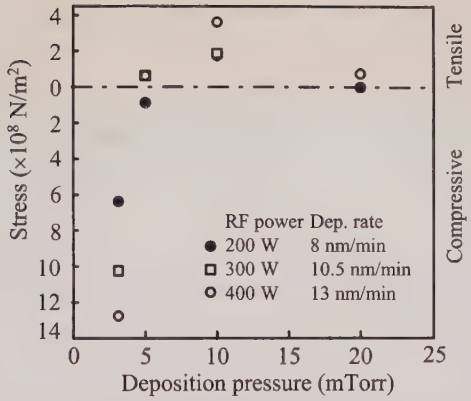


Figure 1.1.6. Measured intrinsic stress in a 500 nm Mo film deposited at 25 °C on intrinsic a-Si:H.

and can exist as SnO and SnO_2. It is well established that the free carriers in ITO come from two principle donors: four-valent Sn^{4+} substituting in the crystalline lattice for In, and doubly charged oxygen vacancies, O^{2+}. Sn acts as an n-type donor which releases electrons to the conduction band in the SnO_2 state. In the SnO state, it does not contribute electrons. The SnO_2 states are predominant in polycrystalline ITO, and the material can be represented as $In_{2-x}Sn_xO_{3-2x}$ [21]. In contrast, the SnO states are predominant in amorphous ITO, hence the lower conductivity associated with the latter.

High crystallinity ITO films have been achieved with high transmittance (>90% in the visible range) and low resistivity ($2\times10^{-4}\,\Omega\cdot$cm) at deposition temperatures over 300 °C [22]. However, room temperature deposited ITO becomes more attractive for the heat-sensitive polymer substrate [23]. The cause of amorphous ITO at low temperatures is the restricted mobility of low kinetic energy indium–oxygen clusters. The low surface mobility preserves the misorientation of the coordination units, which lead to bond distortion. The kinetic energy of adsorbed atoms can be increased by using high sputtering power, low process pressure, and short target–substrate distance in the deposition process. However, the bombardment of the growing film by neutral Ar atoms also increases, which consequently degrades the film integrity. Therefore, the process window for a high crystalline ITO film at low temperatures is narrow. Figure 1.1.7 shows the change in the crystallinity of ITO films at different target–substrate distances [24]. Samples A and B, with target–substrate distances of 12.5 and 13.7 cm, respectively, have high intensities of ⟨222⟩ peaks, which indicates a polycrystalline structure. The wider ⟨222⟩ peak of sample A results from the degradation by high energy Ar bombardment. Samples C and D, with target–substrate distances of 15.4 and 17.7 cm, respectively, show no ⟨222⟩ peaks, which reflects an amorphous structure due to the low kinetic energy of the sputtered atoms on the substrate surface.

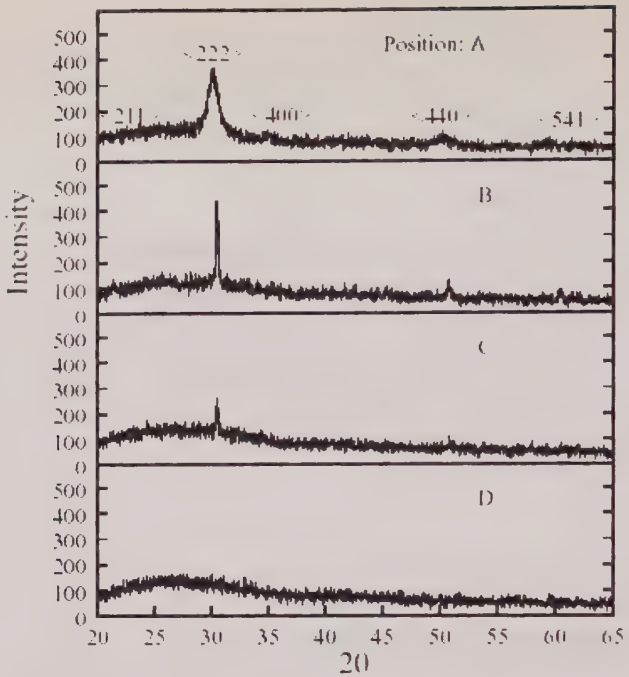


Figure 1.1.7. XRD patterns for ITO films deposited at room temperature with different target/substrate distances. The sputtering rf power is 300 W and process pressure is 15 mT.

The properties of room temperature ITO films depend strongly on the deposition conditions. Polycrystalline ITO films have been achieved with an 80% transmittance in the visible range with a resistivity of $10^{-3}\ \Omega$ cm. The uniformity of the ITO films deposited at room temperature is an important issue. Reasonably good uniformity can be achieved by adding a small amount oxygen with the ar-

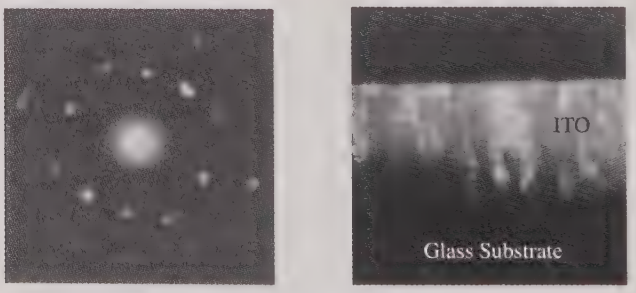


Figure 1.1.8. TEM pictures of polycrystalline ITO films: (**a**) electron diffraction pattern; (**b**) cross section.

gon gas. The electron diffraction pattern of polycrystalline ITO is shown in Figure 1.1.8a. The strong lines of intensities indicate a good crystal orientation. Figure 1.1.8b depicts the columnar structure of the polycrystalline ITO film due to the low surface mobility of the absorbed In and Sn atoms on the growing surface at low temperatures.

1.1.3 Detectors

The choice of detector is largely determined by the detection scheme employed. In direct detection of x-rays, the sensing element is usually a thick photoconductor. In indirect detection of x-rays, the sensing elements most commonly used are based on the Schottky or p-i-n photodiode structures.

1.1.3.1 Photoconductors

In the direct detection approach, a thick photoconductor such as a-Se or lead iodide (PbI_2) absorbs the incident x-ray signal and directly converts this into electrical charge by virtue of the photoelectric effect. Highly energetic photoelectrons are released, which can cause further ionization in the material leading to creation of more e–h pairs. An electric field is applied across the photoconductor to separate the e–h pairs. These carriers traverse through the photoconductor under the electric field and are collected by the electrodes. For full collection, the Schubweg of both holes and electrons should be much longer than the photoconductor thickness. The Schubweg is the mean distance traversed by a carrier before it is trapped, and is given by [5]

$$S = \mu\tau E \tag{1.1.1}$$

where μ is the drift mobility of carriers, τ is the carrier lifetime, and E is the applied electric field. For a-Se, the Schubwegs are $S_e \sim 0.3$ to 3 mm and $S_h \sim 6.5$–65 mm at an applied electric field of 10 V/μm, which are close to the radiological thickness of the a-Se (0.2–1 mm).

In x-ray imaging applications, the a-Se layer is evaporated on a pixel electrode that is connected to a storage capacitor. A continuous electrode is then deposited on top of the a-Se layer to provide the bias contact. Blocking contacts between electrodes and a-Se are incorporated to prevent charge injection from the electrodes into the a-Se [25]. Figure 1.1.9 illustrates the detection schematically. The surface of the a-Se close to the pixel electrode must have a very small transverse conductivity to achieve high spatial resolution. A high density of traps at the a-Se surface is introduced to reduce the surface conductivity. A pure form of a-Se has low thermal stability and tends to crystallize over time. This is not

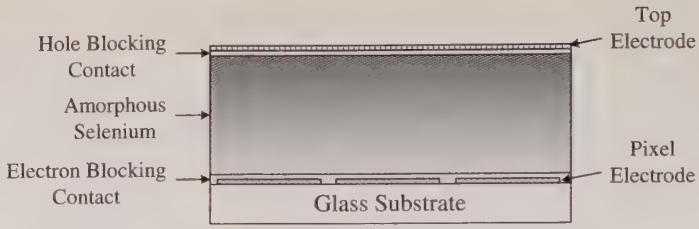


Figure 1.1.9. Schematic diagram of a-Se photoconductor.

suitable for x-ray imaging applications. Alloying the a-Se with about 0.5% arsenic (As) and 10–20 ppm chlorine (Cl) gives a stabilized a-Se with good hole and electron transport [26].

The advantages of a-Se photoconductors are that uniform imaging properties are achieved to a very fine scale, and the material can be easily and cheaply deposited over large areas by a low-temperature process that is compatible with the active matrix array. The disadvantage, however, is the high voltage needed to activate the a-Se layer. Careful design is necessary to avoid the high voltage-induced damage of the active matrix array beneath the a-Se layer [27, 28].

1.1.3.2 Mo/a-Si:H and ITO/a-Si:H Schottky Diodes

A Schottky diode consists of a sandwich structure of metal and a-Si:H (see Figure 1.1.10). The metal could be Mo for direct detection or ITO for indirect detection. The fabrication process of Schottky diodes starts with the sputter deposition of Mo on a glass substrate, which acts as the bottom electrode. Then, n^+a-Si:H, a-Si:H and a-SiN_x:H tri-layers are deposited consecutively in a single pump down at 260°C. The a-Si:H serves as the active layer and the a-SiN_x:H as the passivation layer as well as etch stop layer against the etching of a-Si:H. The final metal (Mo or ITO) is sputter deposited to form the Schottky contact with a-Si:H.

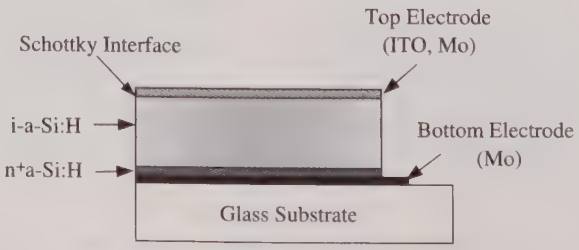


Figure 1.1.10. Schematic diagram of Schottky diode.

Mo/a-Si:H Schottky Diodes

The detection scheme using Mo/a-Si:H Schottky diodes is based on the photoelectric interaction of x-ray photons in the Mo layer and ejection of energetic electrons into the a-Si:H layer [11]. Inside the a-Si:H layer, the energetic electrons undergo various scattering events by which they transfer their energies leading to generation of e–h pairs. The e–h pairs are then separated by the electric field in the depletion region of the reverse biased Schottky diode. The detector quantum yield, ie, the ratio between the measured number of electrons to the number of incident x-ray photons, can be studied by arranging the underlying transduction processes into various stages of conversion as indicated in Figure 1.1.11 [29]. The overall transduction efficiency of the detector (G) for a monoenergetic x-ray beam can be shown in terms of the efficiencies of different stages as

$$G = \eta(\zeta\gamma\chi + \kappa) \tag{1.1.2}$$

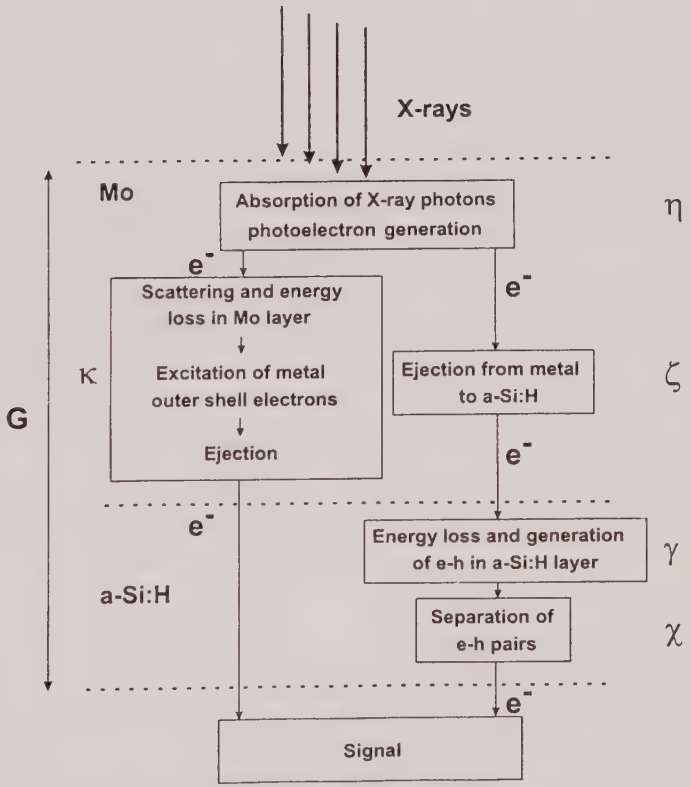


Figure 1.1.11. Schematic representation of the transduction process in the fabricated detectors. Greek symbols refer to the efficiencies of each process.

where η is the photoelectron generation efficiency, ζ the ejection efficiency of the photoelectrons, γ the e–h generation efficiency in the a-Si:H by the ejected photoelectrons, χ the e–h separation efficiency in a-Si:H, and κ the contribution to thermionic emission caused by photoelectrons absorbed in the Mo layer. Equation (1.1.2) is of particular importance in order to examine the impact of different parameters on the detector performance. A quantitative analysis of the various efficiencies in Equation (1.1.2) can be found in [29].

The choice of the top electrode material needs to take into consideration of x-ray absorption, electron ejection, and the intrinsic mechanical stress. The work function of the heavy metal determines the Schottky barrier height, which is crucial in that it determines the detector leakage. The fabricated diodes show rectifying characteristics with an on/off ratio of 10^6 at ± 1 V. The barrier height was measured to be 0.77 eV and the ideality factor was 1.2. The diodes have a low reverse current density (< 10 nA/cm^2 at -1 V); this is crucial for charge retention during imaging. The thickness of the (top electrode) heavy metal is based on a compromise between the absorption of x-rays and the ejection of the energetic electrons from the metal into the depletion layer. Typical x-ray measurement results of a 200 μm^2 Mo/a-Si:H Schottky diode operated at 2 V reverse bias are illustrated in Figure 1.1.12 [29]. These confirm the linear response of the detector to the number of absorbed photons obtained from simulations; the linearity is consistent with data measured for other values of x-ray currents at constant x-ray source voltage. Figure 1.1.13 illustrates the impact of the Mo thickness on the performance of the detector [29]. At low thicknesses, the performance is limited by the absorption of x-ray photons inside the Mo layer. An increase in the thickness of Mo leads to an increase in the number of absorbed photons, and hence the output signal. However, beyond an optimum thickness, the number of electrons reaching the a-Si:H depletion region is limited by the scattering and absorption inside the Mo film. For an x-ray source voltage of 60 kV$_p$, it is found that a thickness of Mo of ~ 500 nm yields the optimum performance. At higher energies, this optimum thickness moves to thicker films. The thickness of the a-

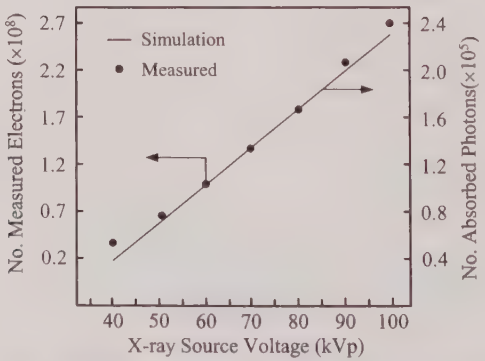


Figure 1.1.12. Response of sensor for various x-ray source voltages (kV$_p$), collected over a period of 500 ms.

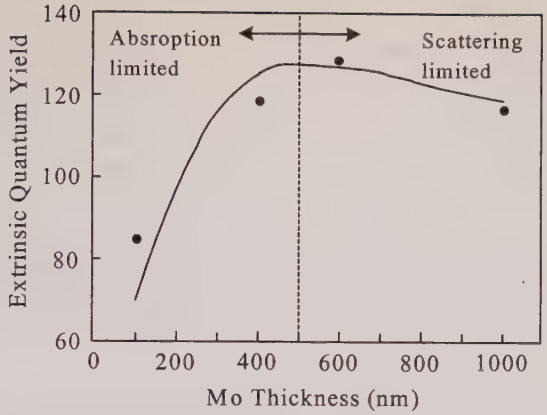


Figure 1.1.13. Variation of sensor response with thickness of Mo layer, taken at 60 kV_p and 25 mA-s.

Si:H layer is another important parameter, which mainly determines the efficiency of stopping the ejected photoelectrons and absorbing their kinetic energies inside a-Si:H. The number of measured electrons increases as the thickness of a-Si:H film increases. However, at large thicknesses the charge collection efficiency at constant bias decreases, since the depletion width remains constant but the neutral region increases. Therefore, the optimum thickness is a compromise between the range of electrons and carrier diffusion length in a-Si:H. Measurement results give a thickness of ~ 1 μm for the optimum thickness for the a-Si:H layer. With respect to bias conditions, the choice is based on a compromise between signal level, noise, reverse current, and reverse current stability [30–32]. Although a low bias results in stable characteristics, low noise, and low leakage current, it yields a lower charge collection efficiency. At higher biases, the charge collection improves, but the stability, noise, and leakage current degrade. Typical values for the noise, stability, detector saturation, and dynamic range, in terms of number of collected electrons over a 500 ms integration period, are 8×10^5, 10^7, 5×10^8, and 50:1, respectively. The use of other metals (eg, W) would yield a larger barrier height with a-Si:H, thus allowing higher voltage operation to enhance charge collection. A larger barrier height would also result in a wider depletion region and reduction in the density of trapped charges to yield an improvement in detector stability.

ITO/a-Si:H Schottky Diodes

In ITO/a-Si:H Schottky diodes, light photons are absorbed in the a-Si:H and the generated e–h pairs are collected by the contacts. When coupled with a phosphor layer for conversion of x-ray to visible photons, ITO/a-Si:H Schottky diodes can be used for x-ray detection. The ITO serves as a light window as well as a Schottky barrier to reduce dark current and minimize noise. The dark current of the Schottky diode, however, is a critical parameter since it limits the resolution and dynamic range of the imaging system.

Owing to the high defect density of a-Si:H, only carriers generated inside the depletion region can be collected. Therefore, Schottky photodiodes are commonly operated under a reverse bias to deplete fully the a-Si:H layer. Electron transport from the ITO to the a-Si:H under reverse bias constitutes an intrinsic leakage current, which limits the dynamic range of the photodiode. There are three transport mechanisms underlying the reverse current in metal/a-Si:H contacts [33, 34]: thermionic emission, thermally assisted tunneling, and field emission. Thermionic emission refers to transport of thermally activated carriers over the top of the Schottky barrier into the a-Si:H conduction band where the conduction band edge intersects with the interface. Thermally assisted tunneling refers to transport of carriers tunneling through the Schottky barrier. Field emission refers to the direct transport of carriers from the Fermi level of the metal to the conduction band of the a-Si:H [35]. Since the defect density is relatively low in the intrinsic a-Si:H layer, the leakage current under medium bias stems mainly from thermionic emission and thermally assisted tunneling. The leakage current density considering both these contributions is expressed as [34, 36, 37]

$$J_0 = J_{s0} \exp(-q\varphi_{beff}/E_0) \exp(V_r/E') \tag{1.1.3}$$

where

$$J_{s0} = \frac{A^*T(\pi qE_{00})^{\frac{1}{2}}}{k} \left\{ q(V_r - E) + \frac{q\varphi_{beff}}{\cosh^2(qE_{00}/kT)} \right\}^{\frac{1}{2}} \tag{1.1.4}$$

$$E_0 = E_{00} \coth(qE_{00}/kT) \tag{1.1.5}$$

$$E' = E_{00}/[qE_{00}/kT - \tanh(qE_{00}/kT)] \tag{1.1.6}$$

$$E_{00} = \hbar(N_d/m^*\varepsilon_r\varepsilon_0)^{\frac{1}{2}}/2 \, . \tag{1.1.7}$$

Here, A^* is the effective Richardson constant, V_r the reverse bias voltage, and φ_{beff} the effective Schottky barrier height. E_{00} can be interpreted as the barrier height when the electron tunneling probability at the edge of the barrier is e^{-1}, and can also be used as an evaluation parameter for the thermally assisted tunneling contribution. If $qE_{00}/kT \ll 1$, E' becomes very large and thermionic emission predominates. If $qE_{00}/kT \sim 1$, then thermally assisted tunneling becomes dominant. Equation (1.1.7) shows that E_{00} increases with ionized defect density, $N_d^{1/2}$, which indicates that thermionic-field emission contribution becomes more significant in high defect density material. J_{s0} is weakly dependent on the bias voltage and equals A^*T^2 when $qE_{00}/kT \ll 1$, and Equation (1.1.3) reduces to the same expression as for thermionic emission.

The primary photocurrent of a thin a-Si:H sample is

$$I_{ph} = \eta_G \eta_c G \tag{1.1.8}$$

where G is the absorbed photon flux, η_G the quantum efficiency and usually equals to unity, and η_c the charge collection efficiency. The charge collection ef-

ficiency depends on the transition of carriers across the depletion region to the electrodes. When carriers travel through the a-Si:H, the defects in the deep states capture the mobile carriers and cause the charge collection efficiency to be less than unity [12]:

$$\eta_c = \frac{Q}{Q_0} = \frac{\mu_D \tau V}{d^2}\left[1 - \exp\left(-\frac{d^2}{\mu_D \tau V}\right)\right] \tag{1.1.9}$$

Here, μ_D is drift mobility, τ the total lifetime for deep trapping, V the applied bias voltage, and d the sample thickness. The condition for full collection is given as

$$d^2 < 0.1\mu_D \tau V_A \approx e a_E V/60\sigma_e N_D kT \approx 2.5 \times 10^7\ V/N_D \tag{1.1.10}$$

where a_E is the minimum scattering length, σ_e the electron capture cross-section, and N_D the density of defect states.

The ITO/a-Si:H Schottky diode has a dark current density of 10^{-8} A/cm^2 and a photocurrent density of 2×10^{-6} A/cm^2 under illumination with a light intensity of 115 μW/cm^2 and at a bias of –2.0 V [36]. Therefore, the dynamic range I_{ph}/I_{dark} is approximately 200. Secondary ion mass spectrometry (SIMS) measurements show that the diffusion of oxygen into a-Si:H is inevitable even when the ITO is deposited at room temperature (see Figure 1.1.14). The oxygen atoms inside the a-Si:H layer act as donor-like defects and increase the density of ionized defect states, which consequently increase the leakage current and reduce the photosensitivity.

The diffusion of oxygen in the Schottky diode is investigated by measurement of the leakage currents of two ITO/a-Si:H photodiodes with room temperature deposited ITO, but with one sample annealed at 260 °C. SIMS measurements show a higher diffusion of oxygen from the ITO to the a-Si:H after annealing [24]. The dark current of the annealed sample is much larger than that of the as-deposited counterpart and increases exponentially with bias voltage, as shown in Figure 1.1.15. The solid lines give the best fit to the reverse current model,

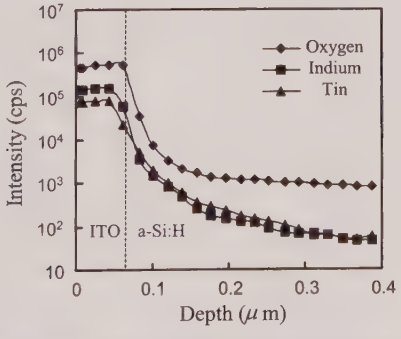


Figure 1.1.14. SIMS measurements of oxygen diffusion in the a-Si:H layer.

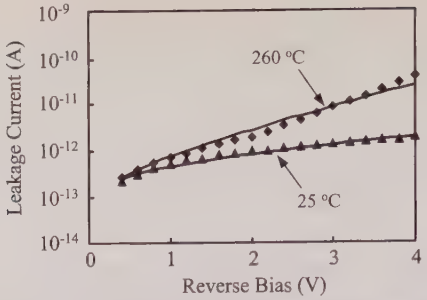


Figure 1.1.15. The reverse currents of ITO/a-Si:H photodiodes for ITO deposited at room temperature and subject to annealing at 260 °C. Solid lines represent best fit to the model (see Equations (1.1.3)–(1.1.7)).

which account for both thermionic emission and thermal assisted tunneling. Table 1.1.3 provides the important parameters extracted from the model [36].

The effective Schottky barrier height under reverse bias is lower than that of the flat band energy owing to presence of the interfacial layer and surface states. The difference is more serious with the annealed sample, which has a high density of defect states in the a-Si:H. The values of the density of surface states, D_s, and the density of defect states, N_{eff}, increase approximately by an order of magnitude after annealing at 260 °C, which is indicative of oxygen diffusion in these devices. An increase of qE_{00}/kT from 0.062 to 0.24 shows the predominance of thermally assisted tunneling after annealing. Therefore, further improvement in the leakage current in ITO/a-Si:H Schottky photodiodes is possible if the oxygen diffusion can be reduced. Photodiodes based on ITO/a-SiN_x:H/a-Si:H MIS structure, with a thin a-SiN_x:H between the ITO and the a-Si:H layers, are proposed to prevent the diffusion of oxygen [38, 39]. A dark current density of 2×10^{-9} A/cm^2 and a dynamic range of 1.4×10^3 are achieved.

Table 1.1.3. Summary of the important parameters extracted from the model (see Equations (1.1.3)–(1.1.7)). R.T. denotes room temperature.

Parameter	Extracted values (R.T.)	Extracted values (260 °C)
m^*/m_0	0.1	0.1
$\varepsilon_i/\varepsilon_0$	3	3
$\varepsilon_s/\varepsilon_0$	11.9	11.9
φ_0 (eV)	0.92	0.92
φ_{off} (eV)	0.86	0.84
δ (nm)	2.8	2.8
D_s (cm^{-2}/eV)	6.2×10^{12}	4.7×10^{13}
N_{deff} (cm^{-3})	7.4×10^{15}	1.36×10^{17}
qE_{00}/kT	0.062	0.24
E'	20	1.4

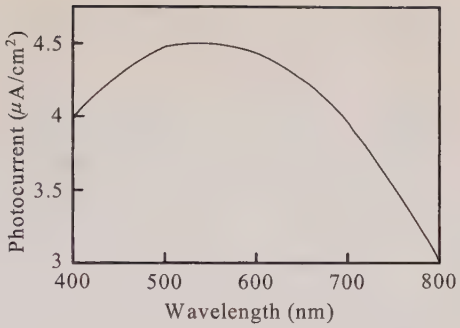


Figure 1.1.16. Spectral response of ITO/a-Si:H Schottky diode.

The spectral response of the photodiode is shown in Figure 1.1.16 for wavelengths in the range 400–800 nm. As expected, the behavior of photocurrent is governed by the absorption characteristics of a-Si:H. The maximum photocurrent is at 570 nm and the spectral response is well matched to the light emitted from the phosphor when subjected to x-rays.

1.1.3.3 p-i-n Photodiodes

In comparison with the Schottky diode, the p-i-n photodiode has a p–i junction in place of the Schottky barrier (see Figure 1.1.17). The p and n layers provide the built-in potential of the junction and block carrier injection from the contacts. The p layer, which can be highly doped a-Si:H or amorphous silicon carbide (a-SiC:H), should be as thin as possible to minimize the absorption of incident photons. The dark current of the p-i-n photodiode is limited by a thermal generation current caused by the thermal excitation of electrons and holes from bulk gap states to the band edges [40, 41]. Beside bulk thermal generation, the additional sources of the dark current are identified to be contact injection, edge leak-

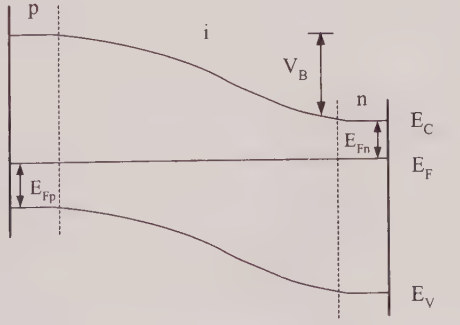


Figure 1.1.17. Band diagram of a p-i-n diode at zero bias.

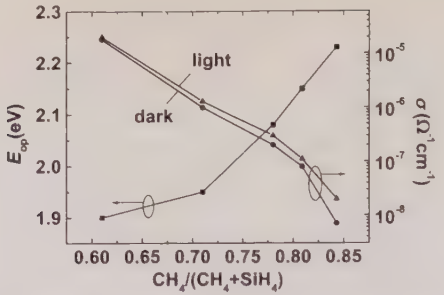


Figure 1.1.18. Effect of methane to silane flow rate on the bandgap and conductivity of the p-type a-SiC:H layer.

age, macrostructural shunt paths, and emission of the carriers from defect states at the p–i and i–n interfaces [42–44].

The a-SiC:H layer is deposited by the dissociation of a gas mixture of SiH_4 and CH_4 in a PECVD chamber, where the ratio of CH_4/SiH_4 flow rates determines the band gap of the material, as shown in Figure 1.1.18. For the p-type a-SiC:H layers, the material is doped by addition of ($\sim$1%) trimethylborane (TMB), which has a lower defect density and higher photoresponse in comparison with samples deposited using diborane (B_2H_6) [44]. In addition, band gap narrowing, which occurs with B_2H_6, is not observed in the TMB case as a function of the CH_4/SiH_4 ratio [45].

Figure 1.1.19 shows cross sections of the different photodiode structures [46–48]. Sample 1 is a p-i-n diode where a p-type a-SiC:H ($\sim$400 Å) is deposited on the Mo-coated glass substrate. Then a thin ($\sim$40 Å) graded layer is deposited by gradually changing the flow rate of CH_4 with respect to SiH_4 followed by an intrinsic a-Si:H layer (1 μm). Then an n-type a-Si:H layer is deposited followed by a thin aluminum ($\sim$200 nm) metallization layer. Sample 2 is a p-δ_i-i-n structure where a thin ($\sim$40 Å) undoped a-SiC:H δ-layer is deposited between the p-type a-SiC:H layer and the graded layer. Sample 3 is an n-i-δ_i-p structure where

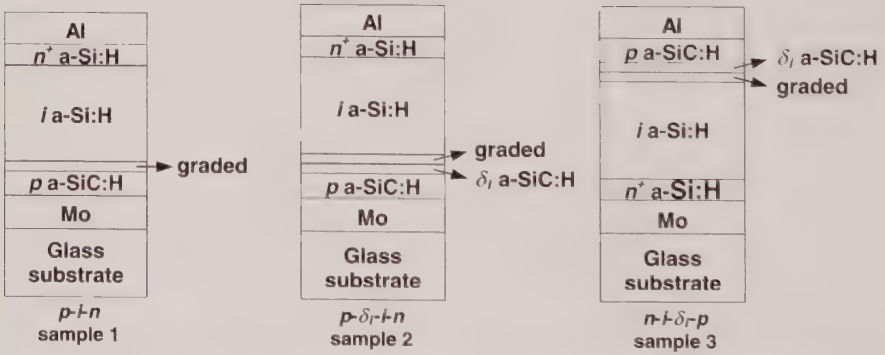


Figure 1.1.19. Cross section of various heterojunction photodiode configurations.

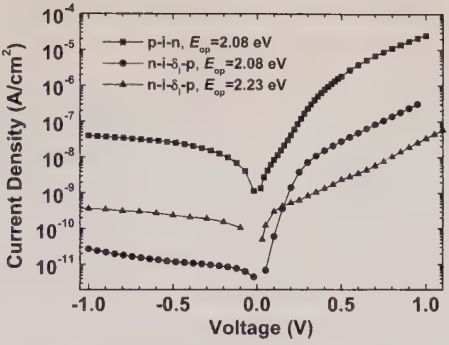


Figure 1.1.20. Dark current–voltage characteristics of the n-i-δ_i-p and the conventional p-i-n a-Si:H/a-SiC:H heterojunction photodiodes with different a-SiC:H bandgap (E_{op}).

the layers of sample 2 are deposited in the reverse order using the same deposition parameters.

The dark current–voltage characteristics at room temperature for the fabricated structures are shown in Figure 1.1.20 [44]. Although all the samples have identical intrinsic a-Si:H layer thickness, the dark currents are distinctly different. The dark current at room temperature is significantly affected by the quality of the p/i interface and cannot be attributed to the thermal generation of e–h pairs through the bulk states. The most probable mechanism responsible for reverse current is the field-enhanced emission from defect states at the interface due to impurities or growth defects [43]. There are two distinct regions of behavior in the forward current–voltage characteristics. In the exponential region at low forward biases ($\leq$0.3 V), the diffusion of electrons and holes from the doped contact layers into the diode bulk dominates, and the current is recombination-limited. In this region, the current–voltage behavior can be described by the diode equation with parameters as given in Table 1.1.4. At higher voltages, the drift current within the diode bulk dominates and the current–voltage behavior becomes non-exponential because the potential profile in the structure changes owing to the trapped charges in the i-layer and the voltage drop across the interface layer.

As can be seen in Figure 1.1.19 and Table 1.1.4, the insertion of the a-SiC:H δ_i-layer significantly reduces the saturation current and reverse current densities, and improves the diode ideality factor. The reduction of the saturation current is

Table 1.1.4. Saturation current density and ideality factor for the different photodiode structures

Configuration	Saturation current density (A/cm^2)	Ideality factor
p-i-n	2×10^{-9}	2.3
p-δ_i-i-n	3×10^{-11}	1.6
n-i-δ_i-p	3×10^{-12}	1.4

about three orders of magnitude, and the ideality factor reduces from 2.3 to 1.4. The high dark current seen in the standard p-i-n photodiode is possibly due to boron-induced defects at the p/i interface [42]. In contrast, for the heterojunction photodiode structures, the low dark current can be attributed to the buffer layer, which serves to reduce boron diffusion. In addition, the average electric field strength in the device structure, and in particular, at the p/i interface reduces, leading to a suppression of the field-enhanced emission of carriers. Thus, the thin undoped a-SiC : H layer acts as an effective blocking layer for reduction of the leakage current. The lowest saturation current density achieved is 3×10^{-12} A/cm^2 and this for the n-i-δ_i-p structure. Although deposited under the same conditions, the saturation current density of the p-δ_i-i–n structure is one order of magnitude higher. Here, the p-layer is subject to the extended processing time (of around 100 min) that is associated with the deposition of the i-layer. Consequently, the buffer layer has an increased density of microstructural defects and is subject to increased boron diffusion, and also higher interface roughness. All of these serve to undermine interface integrity and hence the dark current and ideality factor.

Figure 1.1.21 shows the current–voltage characteristics of the n-i-δ_i-p structure under different illumination intensities. A halogen bulb with a colored glass filter (peak wavelength 520 nm, bandwidth ~ 20 nm) was used as the light source. As can be seen, the collection efficiency depends strongly on the intensity of the incident radiation. The 'S-shape' region of the current–voltage characteristics can be attributed to the presence of the buffer layer. Using the device simulation program AMPS-1D, we analyzed carrier transport in the p-i-n structure with an 8 nm a-SiC : H buffer layer. Figure 1.1.22 shows the internal quantum efficiency of the n-i-δ_i-p diode as a function of wavelength. As can be seen, the recombination at the p/i interface limits the quantum efficiency in the 400–460 nm region [44]. At a wavelength of 550 nm, the quantum efficiency reaches the maximum value of 80% and subsequently degrades at higher wavelengths owing to decreasing absorption.

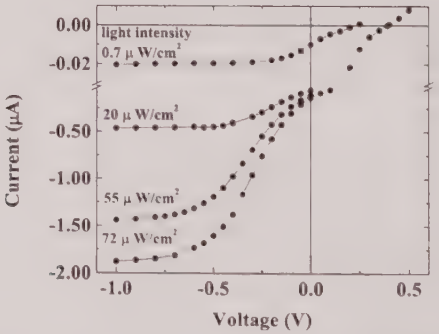


Figure 1.1.21. Current–voltage characteristics of the n-i-δ_i-p sample to different light intensities from a 520 nm light source.

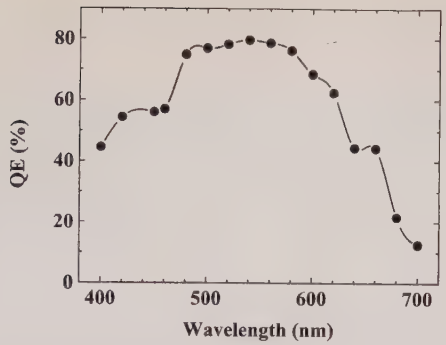


Figure 1.1.22. Internal quantum efficiency of the a-SiC:H/a-Si:H heterojunction photodiode.

1.1.4 Pixel Architectures and Integration

The pixel, forming the fundamental unit of the imager, consists of a detector and readout circuit to transfer the collected electrons efficiently to external readout electronics for data acquisition. The pixel architecture most widely used is based on the passive pixel sensor (PPS). An example is the a-Se-based photoconductor detection scheme where the readout circuit consists of a storage capacitor and readout switch [10]. The storage capacitor accumulates signal charge during the integration period and transfers the collected charge to an external charge amplifier via the TFT switch during readout.

1.1.4.1 Thin Film Transistor and Leakage Current

In AMFPIs, the TFT is used as a switching element in every pixel. The TFTs that have been developed to date include staggered, inverted staggered, coplanar, and dual gate structures [49, 50]. The inverted staggered tri-layer structure (see Figure 1.1.23) is most commonly used in large area imaging applications be-

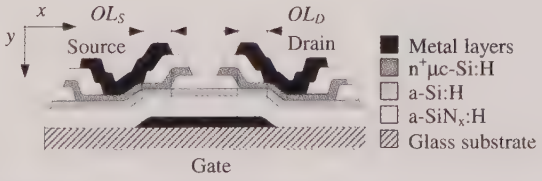


Figure 1.1.23. Cross section of the inverted staggered tri-layer TFT showing its different layers.

cause of the lower interfacial density of states between the gate dielectric and the a-Si:H layer and smoother interface morphology [51, 52].

In conventional imaging arrays, the TFT connects the detector and the storage capacitor (C_S) in the pixel to the read-out circuitry via the data line (Figure 1.1.24). During the integration period, when the TFT is off, the signal charge generated in the detector is stored in C_S. At the same time, the leakage current of the TFT ($I_{leakage}$) and the dark current (I_{dark}) of the detector inevitably mix with the signal. The resulting current at node D obeys the following equation [53]:

$$I_{signal} + I_{dark} - I_{leakage} = C_S \frac{dV_S}{dt} \tag{1.1.11}$$

where I_{signal} is associated with the signal charge and V_S is the voltage of C_S. In high sensitivity imaging applications, where the dark current of the detector is <1 fA, evaluated from 10 pA/cm^2 for a 100 μm^2 detector, a TFT leakage current of the order of 100 fA is the determining source of signal degradation. Hence, depending on the application, the leakage of the TFT becomes a key design constraint, which can limit the performance of the array in terms of its signal-to-noise ratio and dynamic range.

Three mechanisms have been identified as the source of the reverse current: ohmic conduction, front (gate nitride/a-Si:H) channel conduction and back (passivation nitride/a-Si:H) channel conduction [51, 54, 55]. Ohmic conduction takes place across a-Si:H and a-SiN$_x$:H layers according to their intrinsic conductivity. The conduction path for the ohmic current takes place between gate–source, gate–drain, and source–drain. The currents of the first two paths are dependent on the intrinsic conductivity and the quality of both the insulator and a-Si:H layers, which in turn is governed by the fabrication conditions. The third path takes place through the bulk a-Si:H layer between source and drain. The accumulation of holes and electrons at the front and back a-Si:H/a-SiN$_x$:H interfaces, respectively, forms the basis of the other two leakage mechanisms. The relative dominance of the one or the other mechanism depends on bias conditions, TFT geometry, and process conditions. The conduction at the back interface, which prevails at low negative gate voltages, provides the basis of the reverse sub-threshold characteristics of

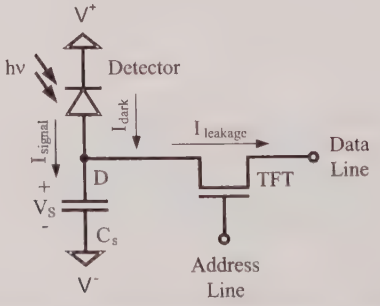


Figure 1.1.24. Circuit schematic of the passive pixel sensor in a-Si:H imaging arrays.

the TFT. The conduction at the front interface prevails at high negative gate voltages. This current component is limited by the Poole-Frenkel electric field-enhanced emission of carriers in the drain overlap area, which reflects the density of the neutral deep traps in the a-Si:H layer [56, 57].

1.1.4.2 Active Pixel Sensors

While the PPS architecture has the advantage of being compact and thus amenable to high-resolution imaging, reading the small output signal of the PPS for low input signal, large area applications (eg, fluoroscopy) is extremely challenging. More important, if external noise sources (eg, charge amplifier noise and array data line thermal noise) are comparable to the input, there is a significant reduction in pixel dynamic range. These problems can be overcome by integrating an on-pixel amplifier circuit using a-Si TFTs [58], along the lines of the CMOS active pixel sensor (APS) architecture [59]. The APS performs in situ signal amplification providing higher immunity to external noise preserving the dynamic range. In addition, the performance and cost constraints on external charge amplifiers are relaxed.

Circuit Operation

Unlike a conventional PPS, which has one TFT switch, there are three TFTs in the APS pixel architecture. This could undermine the fill factor if conventional methods of placing the sensor and TFTs are used. Therefore, in an effort to optimize the fill factor, the TFTs may be embedded under the sensor to provide high fill factor imaging systems. Central to the APS illustrated in Figure 1.1.25 is a

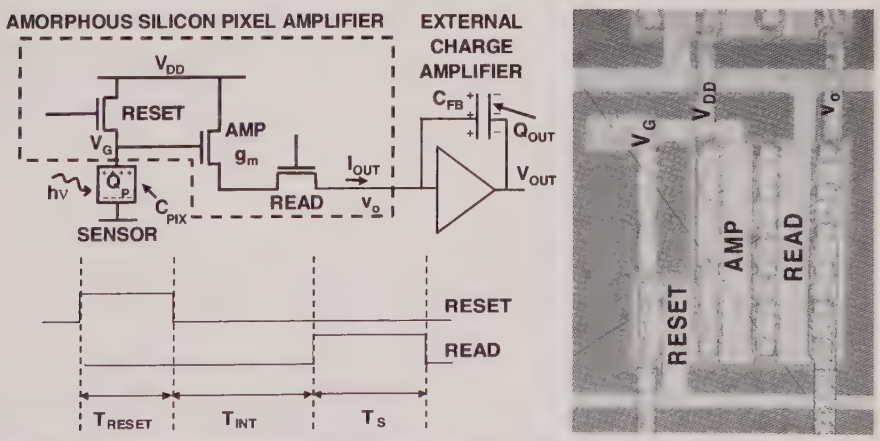


Figure 1.1.25. Current mode active pixel schematic, readout timing diagram and amorphous silicon circuit micrograph (adapted from [58]).

source follower circuit, which produces a current output (C-APS) to drive an external charge amplifier. The C-APS operates in three modes:

- *Reset mode.* The RESET TFT switch is pulsed ON and C_{PIX} charges up to Q_P through the TFT's on resistance. C_{PIX} is usually dominated by the detector (eg, a-Se photoconductor or a-Si photodiode detection layer) capacitance.
- *Integration mode.* After reset, the RESET and READ TFT switches are turned OFF. During the integration period, T_{INT}, the input signal, hv, generates photocarriers discharging C_{PIX} by ΔQ_P and decreases the potential on C_{PIX} by ΔV_G.
- *Readout mode.* After integration, the READ TFT switch is turned ON for a sampling time T_S, which connects the APS pixel to the charge amplifier and an output voltage, V_{OUT}, is developed across C_{FB} proportional to T_S.

In the C-APS circuit, the characteristic ΔV_T of a-Si TFTs is manageable since the TFTs have a duty cycle of $\sim 0.1\%$ in typical large area applications. Therefore, appropriate biasing voltages in the TFT ON and OFF states can minimize ΔV_T. Operating the READ and RESET TFTs in the linear region reduces the effect of inter-pixel threshold voltage (V_T) non-uniformities. However, although the saturated AMP TFT causes the C-APS to suffer from FPN, using CMOS-like off-chip double sampling techniques can alleviate the problem [59].

Linearity

The linearity of the C-APS architecture is obtained from a sensitivity analysis of the change in output current, ΔI_{OUT}, with respect to the input illumination,

$$\gamma = \frac{d(\log|\Delta I_{OUT}|)}{d(\log|hv|)} = \frac{d(\log|\Delta Q_P|)}{d(\log|hv|)} \cdot \frac{d(\log|\Delta V_G|)}{d(\log|\Delta Q_P|)} \cdot \frac{d(\log|\Delta I_{OUT}|)}{d[\log|\Delta V_G|)} , \qquad (1.1.12)$$

where $\gamma = 1$ for an ideal linear sensor, and ΔV_G is the change in the gate voltage of the AMP TFT due to ΔQ_P. The first term is linear if the detector gives a linear change in the charge on C_{PIX}, ΔQ_P with changing hv. The second term depends upon the voltage change ΔV_G across C_{PIX} with changing ΔQ_P, where

$$\Delta Q_P = \Delta V_G C_{PIX} . \qquad (1.1.13)$$

The second term is linear provided that C_{PIX} remains constant under the changing bias conditions. The last term imposes a linear small signal condition on the AMP TFT gate input,

$$\Delta V_G \ll 2(V_G - V_T) \qquad (1.1.14)$$

where V_G is the DC bias voltage at the AMP TFT gate and V_T is its threshold voltage.

Charge Gain

When photons are incident on the detector, e–h pairs are created, leading to a change in the charge given by Equation (1.1.13). In small signal operation, the change in the amplifier's output current with respect to a small change in gate voltage, ΔV_G, is

$$\Delta I_{OUT} = g_m \Delta V_G = g_m V_{in} \tag{1.1.15}$$

where g_m is the transconductance of the AMP and READ TFT composite circuit and V_{in} represents the small signal voltage at the gate of the AMP TFT [60]. Using Equations (1.1.13) and (1.1.15), the charge gain, G_i, stemming from the drain current modulation, is

$$G_i = |\Delta Q_{OUT}/\Delta Q_P| = (\Delta I_{OUT} T_S)/\Delta Q_P = (g_m T_S)/C_{PIX} \tag{1.1.16}$$

The charge gain amplifies the input signal, making it resilient to external noise sources. The theoretical voltage gain and experimental results in Figure 1.1.26 agree reasonably well with a maximum discrepancy of about 20%. Theoretically, using a low capacitance sensor (ie, small C_{PIX}) provides a higher charge gain, which minimizes the effect of external noise. However, a tradeoff between pixel gain and amplifier saturation places an upper limit on the achievable charge gain.

In addition to external noise, investigations of noise added by the APS architecture to the input signal indicate that intrinsic APS noise is minimized for small C_{PIX}, implying the feasibility of low capacitance detectors. In addition, minimizing C_{PIX} will also reduce the reset time constant (which comprises mainly of the RESET TFT on-resistance and C_{PIX}), hence reducing image lag. For example, assuming column parallel readout, a typical array comprising 1000×1000 pixels operating in real-time at 30 frames/s allows 33 μs for each pixel's readout and reset. Typical values for a-Si RESET TFT on resistance (~1 MΩ) and C_{PIX} (~1 pF for a-Se) yield an RC time constant of 1 μs, implying that 5 μs resets would eliminate image lag and still allow sufficient time for readout with double sampling. Like other current mode circuits, the C-APS, operating at 30 kHz, is susceptible to sampling clock jitter. However, off-chip low-jitter clocks using crystal oscillators can alleviate this problem.

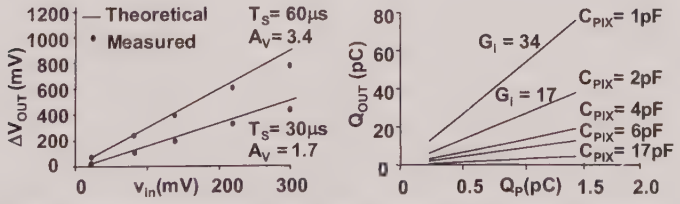


Figure 1.1.26. The measured voltage gain and theoretical charge gain based on different values of C_{PIX}.

1.1.4.3 Pixel Integration

There are number of design requirements that need to be considered for large area x-ray imaging. One of the crucial design requirements in large area imaging is the high (sensor) fill factor, which is defined as the active area expressed as fraction of the physical detector area. In an attempt to achieve a high fill factor, the Schottky diode is stacked on top of the TFT. Two stacked pixel structures (Figure 1.1.27a and b) in the same process are fabricated, and their performances are compared with the conventional pixel structure (Figure 1.1.27c) [61, 62].

In the first structure, the entire TFT is fully overlapped with the diode (Figure 1.1.27a). After the TFT process is completed, the Schottky diode is deposited over the TFT. In the second structure, the overlap is partial. Only the source region of the TFT is overlapped by the Schottky diode (Figure 1.1.27b). In all pixels, the detector is 200×200 μm. Owing to the stacked structure, the channel of the TFT is 200 μm wide and 20 μm long. The stacked pixel structures require 11 masks and 13 lithographic steps. In the nonoverlapping structure (Figure 1.1.27c), which is the standard pixel architecture reported in imaging arrays, the TFT occupies a small portion of the detector area to preserve the high fill factor. The channel of the TFT in the nonoverlapping pixel structure is reduced to 20 μm width and 10 μm length. Since the TFT and diode are not stacked on top of each other, some of the fabrication processes are carried out in parallel. The nonoverlapping pixel structure requires only seven masks and nine lithography steps.

The stacked pixel structures potentially provide high fill factors, but these multi-layer structures give rise to high mechanical stress in the films. The intrinsic stress in these films becomes an important fabrication issue. The optimum thickness of the Mo layer in the Schottky diode is 500 nm. However, when a 500 nm thick Mo layer is deposited on the stacked pixel structure, the layers beneath are unable to withstand the shear stress, particularly at the edge of patterned features, causing the films to crack and peel off the substrate [20], as shown in Figure 1.1.28.

Here, the process sequence starts with the deposition and patterning of 120 nm Cr on a glass substrate to define the TFT gate lines and the bottom electrodes of the photodiodes. Then, 100 nm n^+a-Si:H, 1 μm intrinsic a-Si:H, and 500 nm Mo are sequentially deposited on the Cr patterns. Prior to the next step of patterning the Mo, the film shows cracks and peels off at the edge of Cr patterns (Figure 1.1.29a). A modified process comprising a top-down patterning sequence, as opposed to the original bottom-up process, is proposed to reduce the excessive mechanical stress that concentrates at the sharp corners of the device patterns. Here, Cr, n^+a-Si:H, i-a-Si:H, and Mo films are sequentially deposited. Then the Mo layer is patterned and etched followed by the a-Si:H layer. Finally, Cr is patterned to define the TFT gate lines and the bottom electrodes of the photodiodes. No cracks are observed in this process sequence (Figure 1.1.29b).

Another important requirement in terms of pixel performance is a low leakage current of the TFT. Figure 1.1.30 shows the leakage current in the three different

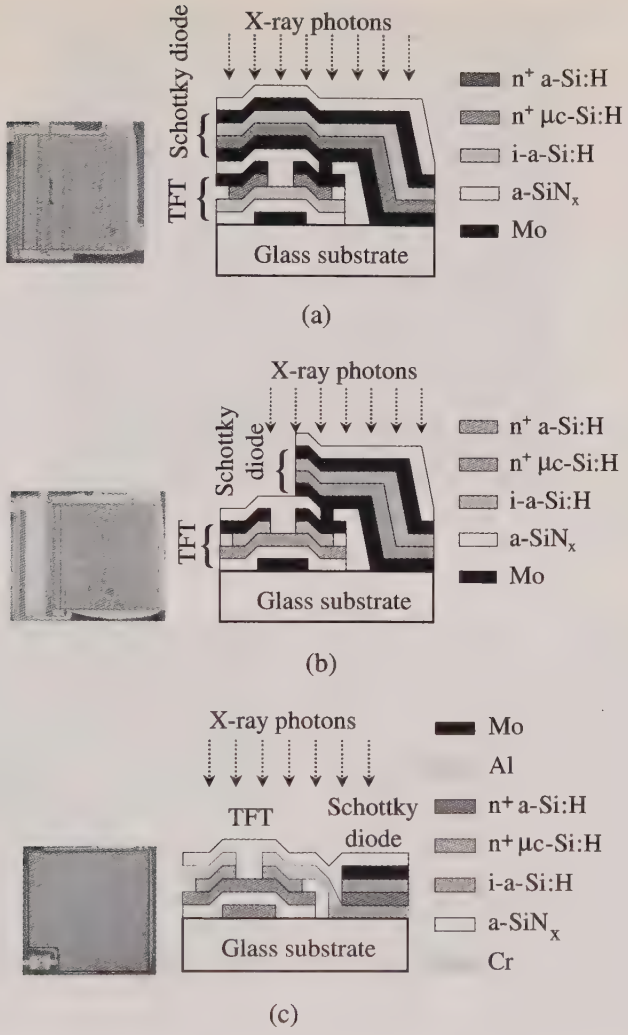


Figure 1.1.27. Cross sectional and top views of different pixel integration structures: (**a**) fully overlapped pixel; (**b**) partial overlapped pixel; and (**c**) nonoverlapping pixel.

pixel structures. In the partially overlapped pixel, the TFT leakage current is as small as that of a discrete TFT ($\sim 10^{-13}$ A). However, the leakage current in the fully overlapped configuration is too high to be employed as a pixel for large area imaging arrays. The high leakage current may be caused by the parasitic back channel, formed at the a-Si:H/a-SiN_x:H interface. In the fully overlapped pixel structure, when the TFT gate voltage is negative, the electrons in the a-Si:H layer experienced an electric field, forming a steady parasitic back channel at the a-Si:H/a-SiN_x:H interface, hence providing a high conducting path from drain to source. This leakage current can be reduced by employing thicker di-

Figure 1.1.28. Top view of damaged thin film layers due to stresses associated with Mo, a-SiN_x:H, and i-a-Si:H layers.

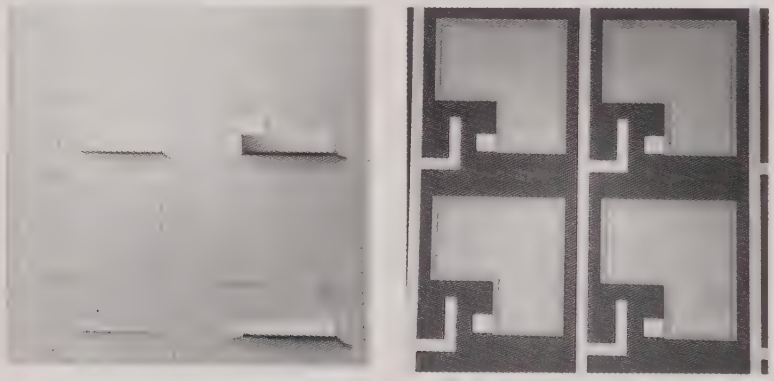


Figure 1.1.29. Photographs of x-ray detectors using (**a**) original and (**b**) modified fabrication processes.

electric layers or by suitably modifying the detector structure to reduce the vertical electric field at the top interface.

The partially overlapped pixel and the nonoverlapping x-ray pixel are exposed to x-rays for 50 ms at various x-ray doses in the range 20–200 mR. The number of measured electrons ranges from 3 to 20 million for the partially overlapped pixel (Figure 1.1.31). However, in the nonoverlapping pixel structure, because of the small TFT geometry, the charges generated by x-rays cannot be as easily transferred to the data line as in the partially overlapped structure with a large TFT, under the same bias conditions. This results in a reduced charge transfer rate and, thus, a reduced number of total measured electrons. As shown in Fig-

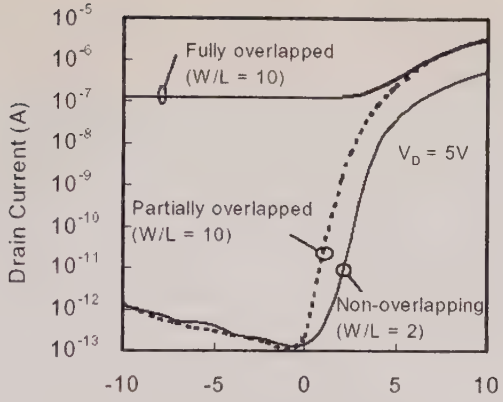


Figure 1.1.30. Leakage currents of TFTs in the different pixel structures.

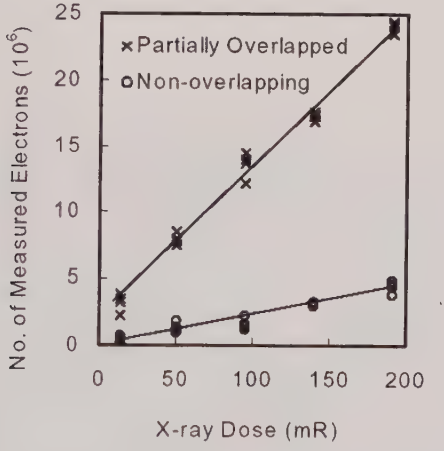


Figure 1.1.31. Number of measured electrons as a function of x-ray dose.

ure 1.1.31, with the same bias conditions as the partially overlapped pixel, the ratio of the measured electrons is approximately 1:5, which is also the ratio of the sizes of the two different TFTs.

1.1.4.4 Phosphor Integration

In indirect x-ray imaging applications, a phosphor film is usually placed on top of the imaging photodiodes. However, the gap between the phosphor film and the sensors can degrade the resolution of the image owing to the light scattering inside the gap. A direct integration of the phosphor film on top of the optical imaging array improves the image quality. Gd_2O_2S:Tb phosphor film is a pow-

der structure combined with small crystal particles (see Figure 1.1.32). Uniformity of the phosphor film over a large area is essential for the integration. The sedimentation technique is used to deposit the phosphor film to provide the large area uniformity for the desired film thicknesses and powder sizes. By adjusting the chemical composition of the coating materials and the operating parameters of sedimentation, a series of phosphor films with $Gd_2O_2S:Tb$ particle sizes from 2.5 to 25 μm and film thicknesses from 85 to 1100 μm can be achieved [63].

The coating solution consists of three major components: $Gd_2O_2S:Tb$, poly(vinyl alcohol) (PVA) and H_2O, coupled with the few organic additives [63], as given in Table 1.1.5. The composite solution provides several controllable parameters, including the concentration of the individual chemicals, the particle size of the phosphor, and the viscosity of the solution. The overall mixing process must provide effective agitation so that the phosphor particles are able to disperse fully into the polymer matrix to achieve a homogeneous coating solution. A high aspect ratio cylinder (wall:diameter=8:1) is implemented to hold the phosphor coating solution for the sedimentation [64]. The settling rates of the sedimentation are proportional to the particle size and inversely proportional to the solution viscosity.

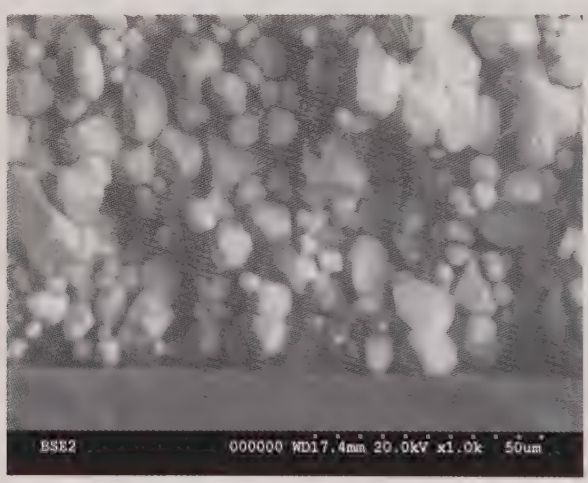


Figure 1.1.32. SEM of $Gd_2O_2S:Tb$ phosphor film.

Table 1.1.5. Chemical components of phosphor coating solution

Component	Units	Concentration
$Gd_2O_2S:Tb$	mg/cm^2	30–200
Poly(vinyl alcohol)	Weight %	3–10
Ethylene glycol based polymer	Weight %	0.2–0.4
Sulfonated fatty acid	Weight %	0.3

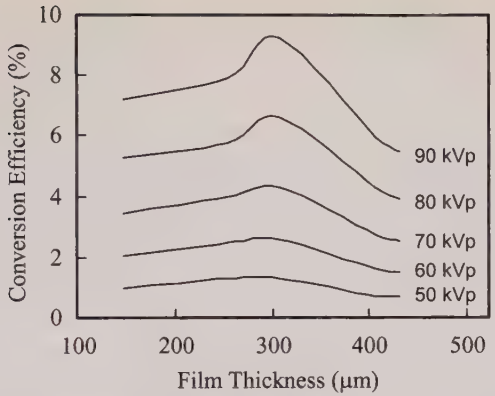


Figure 1.1.33. Conversion efficiency of phosphor films (with a grain size of 2.5 μm) as a function of film thickness for different x-ray energies.

Figure 1.1.33 illustrates the conversion efficiencies of a set of phosphor samples as a function of the film thicknesses with the same particle size under different x-ray energies. A peak conversion efficiency is observed at thickness of 300 μm. The increase in the conversion efficiencies with thickness up to 300 μm is due to the increase in the x-ray absorption efficiency [65]. When the thickness is further increased, the optical photons generated inside the phosphor film travel for a greater distance to reach the film surface, which consequently increases the light loss and decreases the conversion efficiency. Because the Gd_2O_2S:Tb film is composed of powder particles, optical photons generated inside the film isotropically emit, and undergo multiple scattering from the phosphor particles before exiting the film. In thicker films, the photons need to travel longer distances where they are scattered in the process to increase the isotropic propagation path, thereby reducing the spatial resolution. Measurement results of the modulation transfer function (MTF) are illustrated in Figure 1.1.34, which show a degradation in spatial resolution as the phosphor film thickness increases [65].

From the experimental results, thin phosphor films are required to achieve a high spatial resolution. However, the conversion efficiencies of thin phosphor films are low, which reduces the light output from the phosphor film. Thus there is a 'gain–resolution' trade-off with the Gd_2O_2S:Tb phosphor film. In traditional x-ray imaging applications, the imaging array is fully covered by a continuous phosphor film and subject to incident x-rays. The optical photons isotropically traverse the phosphor and reach adjacent photodiodes. The crosstalk is then generated, which degrades the spatial resolution of the imaging array. To avoid this crosstalk, the phosphor film is patterned in such a way that the phosphor material locates only on top of the photodiodes. Therefore, the light generated within a pixel area cannot reach the adjacent pixels, and the crosstalk is minimized. This approach can increase the spatial resolution without sacrificing the gain of the imaging signal. Because the phosphor film has a powder-based structure, and the thickness of the film is more than a few hundred micrometers, it is

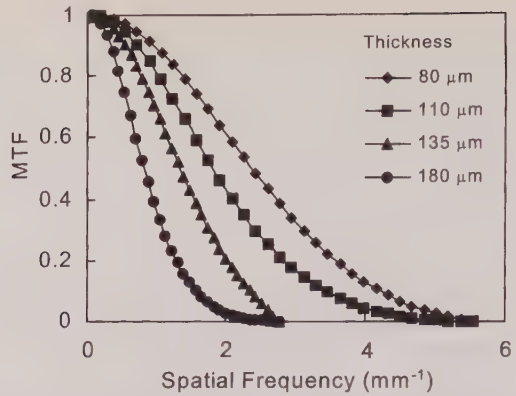


Figure 1.1.34. MTF of phosphor films (with a grain size of 2.5 μm) as a function of spatial frequency for different film thickness.

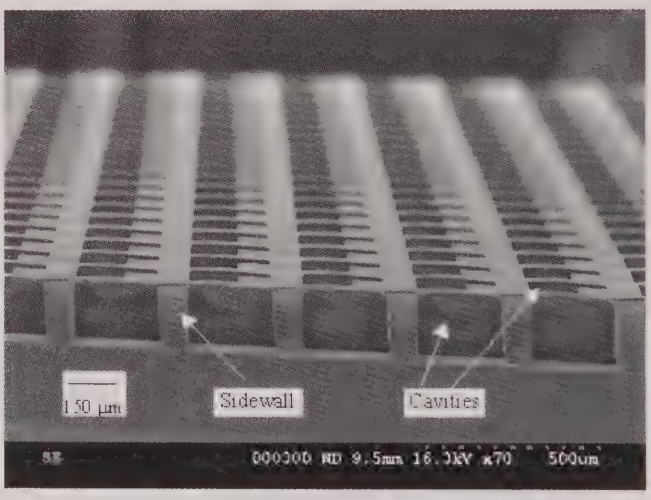


Figure 1.1.35. SEM of patterned SU-8 film of 150 μm thickness.

difficult to pattern the film by traditional IC patterning technologies. A patterning technique using negative photoresist SU-8 [66, 67] was developed to form an array of deep cavities on top of the imaging array (Figure 1.1.35). The phosphor material is then filled inside the cavities by the sedimentation technique to form the patterned phosphor film (Figure 1.1.36). Figure 1.1.37 illustrates the improvement in MTF for a patterned phosphor film compared with the continuous phosphor film [65].

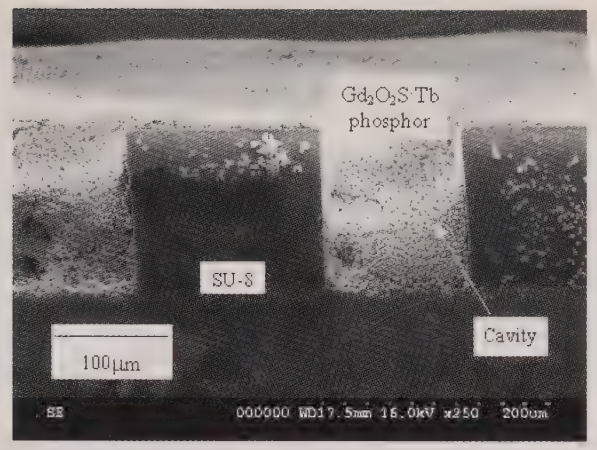


Figure 1.1.36. SEM of a Gd_2O_2S:Tb phosphor film for a grain size 2.5 μm on a patterned SU-8 mould.

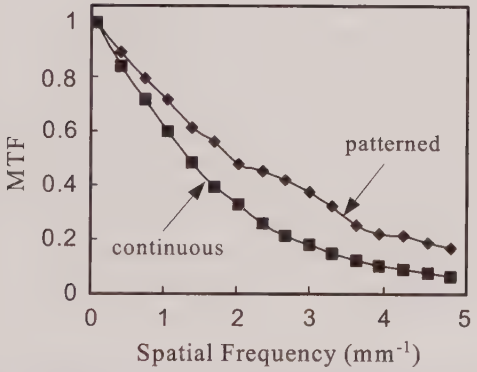


Figure 1.1.37. MTF of continuous and patterned phosphor films for the same film thickness and particle size.

1.1.5 Imaging Arrays

An a-Si:H imaging array consists of a two-dimensional (2-D) matrix of light-sensitive a-Si:H photodiodes that are switched by an active matrix of a-Si:H TFTs [1]. A schematic diagram of the imaging array and its electronics is shown in Figure 1.1.38. Each row of pixels is connected to the gate driver circuitry, which supplies a pulse voltage signal to the gates of the TFTs by horizontal gate lines. Each column of pixels is connected with the drains of the TFTs to external charge integrating amplifiers by vertical data lines. Bias lines connect all the photodiodes to an external bias power supplier.

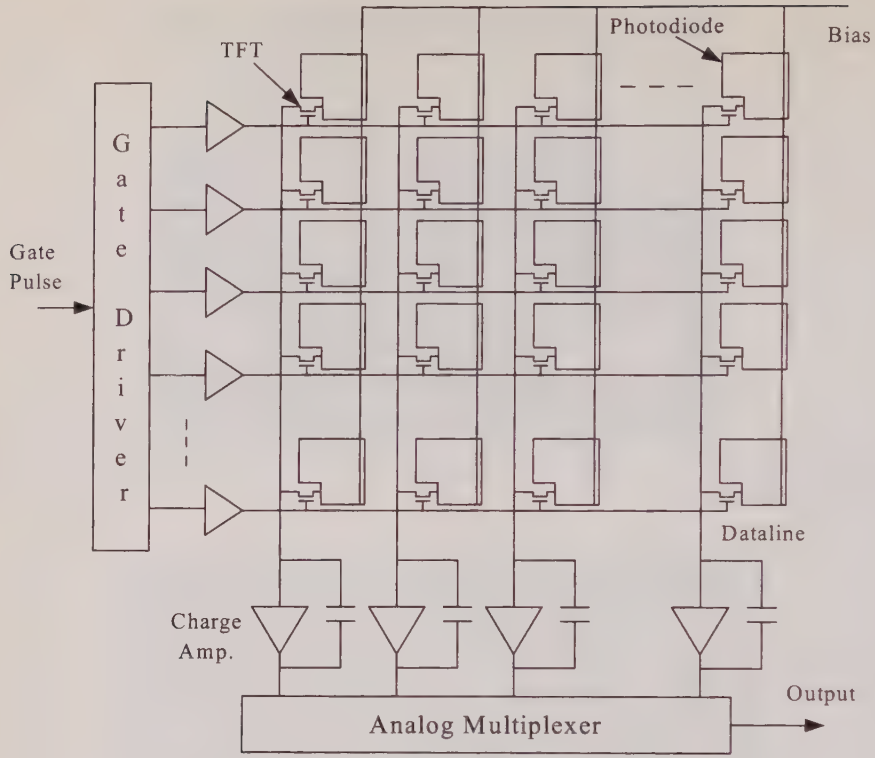


Figure 1.1.38. Schematic diagram of an a-Si:H imaging array.

1.1.5.1 Readout Operation

The photodiodes in the array are operated in the storage mode, which allows the photogenerated charges to be integrated over a long period time to yield a large signal before being read out. The imaging array is read out in the following manner:

- Initially, all the TFTs are in the 'off' state in which the gate voltage V_g is negative so that the TFTs are nonconducting. The photodiodes are reverse biased, and integrate the optical signal during the scanning time.
- The gate driver circuitry then switches the first row of TFTs to the 'on' state by applying a positive gate voltage V_g so that the TFTs along that row are conducting. The integrated signal in the photodiodes from that row propagates along the data lines and appears across the feedback capacitance of the external column charge amplifiers. The signals at the output of the charge amplifiers are then digitized by an A/D converter and sent to a computer for image reconstruction, processing, and display. The first row of the TFTs is held conducting for a duration sufficient to read out the signal in all the photodiodes,

and is then switched back to the 'off' state. Note that the action of reading out the photodiode signal resets the integration node causing the pixel to be simultaneously reset, hence eliminating the need for an additional pixel reset step.

- The next row of pixels is then switched by the gate driver circuitry to the 'on' state and held until the signal from that row is read out, then the gate voltage is switched back to the 'off' state.
- This process is repeated sequentially through each row of the imaging array until all the rows have been read out.

1.1.5.2 Capacitance Extraction for Large Area Arrays

With the growing need for both high pixel density and operating speed, the parasitic coupling capacitances associated with a-Si:H TFTs become a major concern in the realization of large area a-Si:H electronics. The capacitive coupling due to geometric overlap between the gate and source/drain electrodes in the TFT (see Figure 1.1.39) can give rise to charge feedthrough, thus lowering the pixel voltage [68]. To gain an insight into the effect of parasitic coupling capacitances on the overall array performance, these capacitances need to be extracted in an efficient manner and with a high degree of accuracy. A common analytical tool for capacitance estimation is the parallel-plate formula. However, the suitability of this simple formula is restricted because it relies only on the overlap area and does not take into account the effect of the fringing field or the dielectric media surrounding the conductors. For example, with the parallel-plate formula, a zero overlap area would yield zero capacitance, which is incorrect since there can be nontrivial capacitive coupling due to the fringing field. Therefore, numerical simulation is imperative. From a computational standpoint, the main difficulties in the numerical extraction of capacitance include:

- the extreme device geometries, in which the ratio of thin film thickness to other physical dimensions (such as the width and length) can well be of the order of 10^{-3};

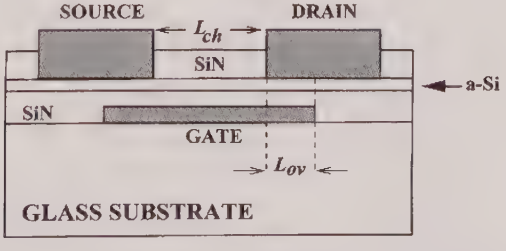


Figure 1.1.39. Schematic cross section of an a-Si:H TFT indicating source–gate and drain–gate overlaps.

- the floating potential of the glass substrate: the substrate now needs to be included as part of the computational domain;
- treatment of multi-dielectric media where the electric field is discontinuous across dielectric interfaces.

The extreme geometry of metal lines and active devices and the presence of the glass substrate require a large number of panels (or mesh elements) barely to resolve the surfaces of the conductors and the interfaces between two dielectric media. This leads to a very large system of equations. Furthermore, treatment of multi-dielectric media necessitates more accurate approximation of the potential.

These difficulties can be problematic and may even be out of reach for numerical methods, which impose exorbitant memory requirements. However, these concerns can be addressed effectively using the exponential expansion (EE) based method, which has been developed recently for efficient evaluation of the three-dimensional potential field [69].

Among the main features of the EE method are: (i) the memory requirement is independent of the desired degree of accuracy on the potential and (ii) different forms of parallelisms are available for both remotely distributed networks and closely coupled parallel systems. These features make it possible for accurate large-scale simulation for capacitance extraction in TFTs involving panel counts in the range from a few hundred thousand to several million. A comparison of the EE method with other well known schemes based on multipole expansion is given in Table 1.1.6. The geometric overlap capacitance between the gate and source/drain electrodes of an a-Si:H TFT is computed based on a panel count for discretization of conductor surfaces and dielectric interfaces of nearly 400000. The length, width, and thickness of the structure considered (see Figure 1.1.39) in simulations are $100\times100\times11.3$ μm. Here, only 10 μm of the glass substrate ($\varepsilon_r=5.84$) (or silicon substrate ($\varepsilon_r=11.9$)) is considered so as to reduce the panel count. The respective layer thicknesses and dielectric constants are listed in Table 1.1.7. The length of the source and drain region is 25 μm while that of the channel (L_{ch}) is 15 μm. The gate length is varied from 10 to 32 μm, which corresponds to an overlap length (L_{ov}) from –2 to 9 μm. Here, the nega-

Table 1.1.6. Multipole expansion (MP) versus exponential expansion (EE) methods (p and S_{app} denote the order of expansion)

	Multipole expansion	Exponential expansion
Method	Fast multipole algorithm (FMA) [70]	EE method
Accuracy	$1/2^p$	10^{-2}–10^{-16}
Time	$O(p^3N)$	$O(S_{app}N)$
Memory	$O(p^2N)$	$O(N)$
Applications	FASTCAP [71] BEMMODULE [72]	This work

Table 1.1.7. Thicknesses of various layers and their dielectric constants

	a-SiN_x (gate)	a-Si:H	a-SiN_x (top)	Source and drain metal	Gate metal
Thickness (μm)	0.26	0.05	0.25	1.0	0.12
ε_r	4.5	11.9	3.1		

Table 1.1.8. Geometric overlap capacitance form parallel plate approximation and numerical simulation

L_{ov} (μm)	C_{approx} (fF)	$C_{simulation}$ (fF)	
		Glass (ε_r=5.84)	Si (ε_r=11.9)
−2	0	6.52	11.0
−1	0	8.55	14.5
0	0	16.1	23.0
1	26.7	42.7	49.3
2	53.5	70.6	76.3
3	80.3	97.2	103
6	170	189	192
9	241	260	262

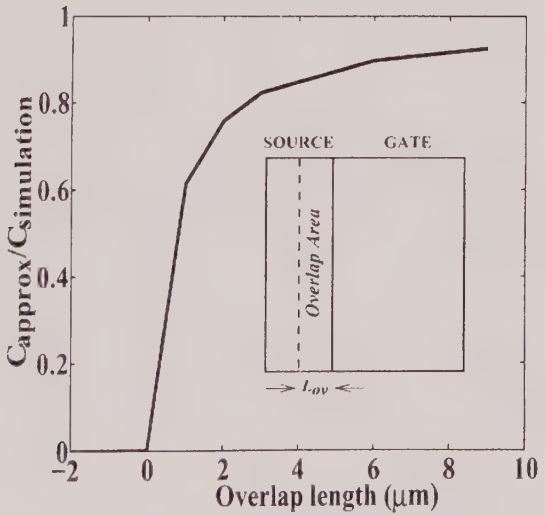


Figure 1.1.40. Fringing factor $\alpha = C_{approx}/C_{simulation}$ as a function of overlap length L_{ov}. The substrate is glass.

tive sign implies no overlap and the associated value denotes the horizontal separation between the gate and the source/drain electrodes.

The values of the geometric overlap capacitance between gate and source are listed in Table 1.1.8 for different L_{ov}. These values are obtained from parallel-plate approximation (column 2) and numerical simulation for glass substrate (column 3) and silicon substrate (column 4). From the simulations, it is observed that (i) the capacitance is a nonlinear function of L_{ov}, and hence the overlap area (see Figure 1.1.40); (ii) even when $L_{ov} \leq 0$, the capacitance is nontrivial; and (iii) the capacitance increases with increasing permittivity of the substrate material (see Table 1.1.8).

1.1.6 New Challenges in Large Area Digital Imaging

One of the key issues with AMFPIs lies in reduction of pixel size [73]. This may be achieved using the approaches of stacked pixel architectures discussed in Section 1.1.4.3, which provide 100% fill factors. However, the challenges here lie in reducing the pixel's parasitic capacitance, leakage current, and mechanical stress associated with the multiple thin film stack. Alternatively the pixel size in the nonoverlapping pixel architecture can also be reduced, without compromising the fill factor, through use of a vertical TFT structure (VTFT) for the active matrix switch [74, 75]. In contrast to the conventional lateral structure, the channel length in the VTFT is defined by the a-SiN_x:H layer thickness in a drain-source island, and thus the channel length can be precisely controlled by film thickness. Crucial benefits of this device are not only in high fill factor imaging, but also for other high resolution applications, including displays.

Recently, there has been growing interest in the fabrication of large-area TFT arrays on thin flexible plastic substrates. This stems from the need for low-cost, light-weight, unbreakable, and foldable computer screens or large-area sensor arrays for imaging of nonplanar (curved) surfaces. However, there are several technological issues that need to be addressed, that range from reducing the deposition temperature of a-Si:H and a-SiN_x:H to optimizing the mechanical stress in these layers, and also in the composite device structure, for compatibility with plastic substrates. Preliminary results of fabrication on a Kapton substrate of the active matrix switches are very encouraging [76]. Devices with OFF currents $< 10^{-12}$ A, ON currents $> 10^{-5}$ A, and device mobility close to 1 cm^2/V s have been achieved [77, 78].

1.1.7 References

[1] Street, R.A., Wu, X.D., Weisfield, R., Ready, S., Apte, R., Nguyen, M., Nylen, P., *Mater. Res. Soc. Symp. Proc.* **377** (1995) 757–766.
[2] Rowlands, J.A., Zhao, W., Blevis, I.M., Waechter, D.F., Huang, Z., *RadioGraphics* **17** (1997) 753–760.
[3] Weisfield, R.L., Street, R.A., Apte, R., Moore, A., *Proc. SPIE* **3032** (1997) 14–21.
[4] Nathan, A., Park, B., Sazonov, A., Tao, S., Chan, I., Servati, P., Karim, K., Charania, T., Striakhilev, D., Ma, Q., Murthy, R.V.R., *Microelectron. J.* **31** (2000) 883–891.
[5] Beutel, J., Kundel, H.L., Van Metter, R.L. (eds.) *Handbook of Medical Imaging, Vol. 1, Physics and Psychophysics*; Bellingham, WA: SPIE, 2000; www.spie.org/bookstore/.
[6] van Berkel, C., Powell, M.J., Deane, S.C., *J. Non-Cryst. Solids* **164–166** (1993) 653–658.
[7] LeComber, P.G., Spear, W.E., Ghaith, A., *Electron. Lett.* **15** (1979) 179–181.
[8] Murthy, R.V.R., Ma, Q., Nathan, A., Chamberlain, S.G., *Mater. Res. Soc. Symp. Proc.* **507** (1998) 73–78.
[9] Antonuk, L.E., El-Mohri, Y., Huang, W., Siewerdsen, J.H., Yorkston, J., Street, R.A., *Mater. Res. Soc. Symp. Proc.* **336** (1994) 855–860.
[10] Zhao, W., Rowlands, J.A., *Med. Phys.* **22** (1995) 1595–1604.
[11] Aflatooni, K., Nathan, A., Hornsey, R., *Tech. Dig. IEEE IEDM* (1997) 197–200.
[12] Street, R.A., *Hydrogenated Amorphous Silicon*; Cambridge: Cambridge University Press, 1991.
[13] Staebler, D.L., Wronski, C.R., *Appl. Phys. Lett.* **31** (1977) 292–294.
[14] Maeda, K., Umezu, I., *J. Non-Cryst. Solids* **227–230** (1998) 43–47.
[15] Kamei, T., Matsuda, A., *J. Vac. Sci. Technol. A* **17** (1999) 113–120.
[16] Morimoto, A., Matsumoto, M., Yoshita, M., Kumeda, M., *Appl. Phys. Lett.* **59** (1991) 2130–2132.
[17] Masuda, A., Itoh, K.-I., Zhou, J.H., Kumeda, M., Shimizu, T., *Jpn. J. Appl. Phys.* **33** (1994) L1295–L1297.
[18] Park, B., Murthy, R.V.R., Benaissa, K., Aflatooni, K., Nathan, A., Hornsey, R.I., Chamberlain, S.G., *J. Vac. Sci. Technol. A* **16** (1998) 902–905.
[19] Robertson, J., Powell, M., *Mater. Res. Soc. Symp. Proc.* **49** (1985) 215–222.
[20] Park, B., Karim, K.S., Nathan, A., *J. Vac. Sci. Technol. A* **18** (2000) 688–692.
[21] Paine, D.C., Whitson, T., Janiac, D., Beresford, R., Yang, C.O., *J. Appl. Phys.* **85** (1983) 8445–8450.
[22] Ray, S., Banerjee, R., Basu, N., Batabyal, A.K., Barua, A.K., *J. Appl. Phys.* **54** (1983) 3497–3501.
[23] Zhu, F., Zhang, K., Guenther, E., Jin, C.S., *Thin Solid Films* **363** (2000) 314–317.
[24] Ma, Q., Nathan, A., *Electrochem. Solid-State Lett.* **2** (1999) 408–420.
[25] Johanson, R.E., Kasap, S.O., Rowlands, J.A., Polischuk, B., *J. Non-Cryst. Solids* **227–230** (1998) 1359–1362.
[26] Shukri, Z., Caron, M., Rougeot, H., *Proc. SPIE* **3659** (1999) 48–55.
[27] Den Boer, W., Aggas, A., Byun, Y.H., Gu, T., Zhong, J.Z.Z., Thomsen, S.V., Jeromin, L.S., Lee, D.L.Y., *Proc. SPIE* **3336** (1998) 520–528.
[28] Zhao, W., Law, J., Waechter, D., Huang, Z., Rowlands, J.A., *Med. Phys.* **25** (1998) 539–549.

[29] Nathan, A., Hornsey, R., Aflatooni, K., *IEEE Trans. Electron Devices* **47** (2000) 2093–2100.
[30] Hornsey, R.I., Aflatooni, K., Nathan, A., *Appl. Phys. Lett.* **70** (1997) 3260–3262.
[31] Aflatooni, K., Hornsey, R., Nathan, A., *IEEE Trans. Electron Devices* **46** (1999) 1417–1422.
[32] Aflatooni, K., Nathan, A., Hornsey, R.I., *Mater. Res. Soc. Symp. Proc.* **420** (1996) 747–752.
[33] Sze, S.M., *Physics of Semiconductor Devices*; New York: Wiley, 1981.
[34] Rhoderick, E.H., *Metal–Semiconductor Contacts*; Oxford: Oxford University Press, 1978.
[35] Crowell, C.R., Rideout, V.L., *Solid-State Electron.* **12** (1969) 89–105.
[36] Tao, S., Ma, Q., Striakhilev, D.S., Nathan A., *Mater. Res. Soc. Symp. Proc.* **609** (2001) A12.2.1–6.
[37] Kramer, N., van Berkel, C., *Appl. Phys. Lett.* **64** (1994) 1129–1131.
[38] Vink, T.J., Nieuwesteeg, J.B.M., Oversluizen, G., *J. Appl. Phys.* **71** (1992) 4399–4404.
[39] Rothwarf, A., Pereyra, I., *Solid-State Electron.* **24** (1981) 1067–1070.
[40] Street, R.A., *Philos. Mag. B* **63** (1991) 1343–1363.
[41] Schiff, E.A., Street, R.A., Weisfield, R.L., *J. Non-Cryst. Solids* **198–200** (1996) 1155–1158.
[42] McMahon, T.J., Yacobi, B.G., Sadlon, K., Dick, J., Madan, A., *J. Non-Cryst. Solids* **66** (1984) 375–380.
[43] Arch, J.K., Fonash, S.J., *J. Appl. Phys.* **72** (1992) 4483–4485.
[44] Morrison, S., Servati, P., Vygranenko, Y., Nathan, A., Madan, A., *Mater. Res. Soc. Symp. Proc.* **715** (2002) A7.4.1–A7.4.6.
[45] Kuwano, Y., Tsuda, S., *Springer Proc. Phys.* **34** (1989) 167–177.
[46] Schropp, R.E.I., et al., *Mater. Res. Soc. Symp. Proc.* **297** (1993) 797–802.
[47] Arya, R.R., Catalono, A., Oswald, R.S., *Appl. Phys. Lett.* **49** (1986) 1089–1091.
[48] Kim, W.Y., Tasaki, H., Konagai, M., Takahashi, K., *J. Appl. Phys.* **61** (1987) 3071–3076.
[49] Kuo, Y., *Proc. SPIE* **1117** (1989) 114–122.
[50] Kuo, Y., *J. Electrochem. Soc.* **142**, 2486–2507.
[51] Street, R.A., Thompson, M.J., *Appl. Phys. Lett.* **45** (1984) 769–771.
[52] Chikamura, T., Hotta, S., Nagata, S., *Mater. Res. Soc. Symp. Proc.* **95** (1987) 421–430.
[53] Servati, P., Nathan, A., *IEEE Trans. Electron Devices* **49** (2002) 812–819.
[54] Shannon, J.M., Morgan, B.A., *J. Appl. Phys.* **86** (1999) 1548–1551.
[55] Street, R.A., Thompson, M.J., Johnson, N.M., *Philos. Mag. B* **51** (1985) 1–17.
[56] Servati, P., Nathan, A., *J. Vac. Sci. Technol. A* **20** (2002) 1038–1042.
[57] Servati, P., Nathan, A., Sazonov, A., *Mater. Res. Soc. Symp. Proc.* **609** (2000) A28.3.1–6.
[58] Karim, K.S., Nathan, A., *IEEE Electron Device Lett.* **22** (2001) 469–471.
[59] Mendis, S.K., Kemeny, S.E., Fossum, E.R., *IEEE Trans. Electron Devices* **41** (1994) 452–453.
[60] Matsuura, N., Zhao, W., Huang, Z., Rowlands, J.A., *Med. Phys.* **26** (1999) 672–681.
[61] Park, B., Murthy, R.V.R., Sazonov, A., Nathan, A., Chamberlain, S.G., *Mater. Res. Soc. Symp. Proc.* **507** (1999) 237–242.
[62] Park, B., Nathan, A., *Proc. ECS* **98-22** (1999) 381–390.
[63] Gu, Z.H., Tao, S., Chan, I., Nathan, A., *J. Vac. Sci. Technol. A* **18** (2000) 639–642.
[64] Rhodes, M., *Introduction to Particle Technology*; New York: Wiley, 1998.

[65] Tao, S., Gu, Z.H., Nathan, A., *J. Vac. Sci. Technol. A* **20** (2002) 1091–1094.
[66] Shaw, J.M., Geloreme, J.D., LaBianca, N.C., Conley, W.E., Holmes, S.J., *IBM J. Res. Dev.* **41** (1997) 81–94.
[67] Zhang, J., Tan, K.L., Hong, G.D., Yang, L.J., Gong, H.Q., *J. Micromech. Microeng.* **11** (2001) 22–26.
[68] Tsukada, T., *J. Non-Cryst. Solids* **164–166** (1993) 721–726.
[69] Pham, H., Nathan, A., *Proc. R. Soc. London, Ser. A* **455** (1999) 637–675.
[70] Greengard, L., *The Rapid Evaluation of Potential Fields in Particle Systems*; Cambridge, MA: MIT Press, 1988.
[71] Nabors, K., Kim, S., White, J., *IEEE Trans. Microwave Theory Tech.* **40** (1992) 1496–1506.
[72] Bächtold, M., Korvink, J., Baltes, H., *IEEE Trans. Comput.-Aided Des.* **15** (1996) 1541–1546.
[73] Nathan, A., Park, B., Ma, Q., Sazonov, A., Rowlands, J.A., *Microelectron. Reliab.* **42** (2002) 735–746.
[74] Uchida, Y., Nara, Y., Matsumura, M., *IEEE Electron Device Lett.* **EDL-5** (1984) 105–107.
[75] Chan, I., Nathan, A., *J. Vac. Sci. Technol. A* **20** (2002) 962–965.
[76] Gleskova, H., Wagner, S., Suo, Z., *Mater. Res. Soc. Symp. Proc.* **508** (1998) 73–78.
[77] Sazonov, A., Nathan, A., Murthy, R.V.R., Chamberlain, S.G., *Mater. Res. Soc. Symp. Proc.* **558** (1999) 375–380.
[78] Stryahilev, D., Sazonov, A., Nathan, A., *J. Vac. Sci. Technol. A* **20** (2002) 1087–1090.

List of Symbols and Abbreviations

Symbol	Designation
A^*	effective Richardson constant
C_S	storage capacitance
d	sample thickness
D_S	density of surface states
E	electric field
G	absorbed photon flux
G_i	charge gain
g_m	transconductance of composite circuit
I_{dark}	dark current of detector
$I_{leakage}$	leakage current of TFT
I_{OUT}	output current
I_{signal}	current associated with signal charge
J_0	leakage current density
L_{ch}	channel length
L_{ov}	overlap length
N_D	density of defect states

Symbol	Designation
S	Schubweg
T_{INT}	integration period
V	applied bias voltage
V_G	gate voltage
V_{OUT}	output voltage
V_r	reverse bias voltage
V_S	voltage of storage capacitor
V_T	threshold voltage
a_e	minimum scattering length
χ	electron–hole separation efficiency
ε	dielectric constant
η	photoelectron generation energy
η_c	charge collection efficiency
η_G	quantum efficiency
φ_{eff}	effective Schottky barrier height
γ	electron–hole generation efficiency
κ	photoelectron contribution to thermionic emission
μ	drift mobility of carrier
μ_D	drift mobility
σ_e	electron capture cross section
τ	total lifetime for deep trapping
ζ	ejection efficiency of photoelectrons

Abbreviation	Explanation
AMFPI	active matrix flat panel imager
APS	active pixel sensor
a-Se	amorphous selenium
a-Si	amorphous silicon
CCD	charge-coupled device
CMOS	complementary metal oxide semiconductor
c-Si	crystalline silicon
EE	exponential expansion
e–h	electron–hole
IC	integrated circuit
ITO	tin-doped indium oxide
PECVD	plasma enhanced chemical vapor deposition
PPS	passive pixel sensor
PVA	polyvinyl alcohol
PVD	physical vapor deposition
SIMS	secondary ion mass spectrometry

Abbreviation	Explanation
TEM	transmission electron microscope
TFT	thin film transistor
TMB	trimethylborane
UHV	ultra-high vacuum
VTFT	vertical thin film transistor

1.2 CMOS Single-chip Gas Detection Systems: Part II

C. Hagleitner, A. Hierlemann and H. Baltes,
Physical Electronics Lab, ETH Zurich, Zurich, Switzerland

Abstract

Sensor arrays based on industrial CMOS-technology combined with post-CMOS micromachining (CMOS MEMS) are a promising approach to low-cost sensors. In the first part of this article [1], the state of research on CMOS-based gas sensor systems was reviewed, and a platform technology for monolithic integration of three different transducers on a single chip was described. In this second part, the transduction principles of three polymer-based gas sensors are detailed and the read-out circuitry is portrayed. The first transducer is a micromachined resonant cantilever. The absorption of analyte in the chemically sensitive polymer causes shifts in resonance frequency as a consequence of changes in the oscillating mass. The cantilever acts as the frequency-determining element in an oscillator circuit, and the resulting frequency change is read out by an on-chip counter. The second transducer is a planar capacitor with polymer-coated interdigitated electrodes. This transducer monitors changes in the dielectric constant upon absorption of the analyte into the polymer matrix. The sensor response is read out as a differential signal between the coated sensing capacitor and a passivated reference capacitor, both of which are incorporated into the input stage of a switched capacitor second-order ΣΔ-modulator. The third transducer is a thermoelectric calorimeter, which detects enthalpy changes upon ab-/desorption of analyte molecules into a polymer film located on a thermally insulated membrane. The enthalpy changes in the polymer film cause transient temperature variations, which are detected via polysilicon/aluminum thermocouples (Seebeck effect). The small signals in the μV-range are first amplified with a low-noise chopper amplifier, then converted to a digital signal using a ΣΔ-A/D-converter and finally decimated and filtered with a digital decimation filter.

Keywords: chemical sensors, monolithic gas sensor, integrated sensor, interface circuitry, polymer, chopper amplifier, low-noise amplifier, differential difference amplifier, ΣΔ-modulator, decimation filter, cantilever, calorimeter

Contents

1.2.1 Polymers as Sensitive Layers . . . 53
1.2.1.1 Polymers for the Detection of VOCs . . . 53
1.2.1.2 Polymer–Analyte Interaction . . . 54
1.2.1.3 Modeling of the Polymer–Analyte Interaction . . . 55

1.2.2 Capacitive Sensor . . . 57
1.2.2.1 Sensing Principle . . . 57
1.2.2.2 Equivalent Circuit Model . . . 58
1.2.2.3 Read-out Circuitry . . . 59
1.2.2.4 Second-order Sigma–Delta Modulator . . . 60
1.2.2.4.1 Single-bit-modulation Techniques for Sensor Interfaces . . . 60
1.2.2.4.2 Second-order Sigma–Delta Modulator . . . 61
1.2.2.4.3 Tunable Reference Capacitor . . . 63
1.2.2.4.4 Interdigitated Feedback Capacitors . . . 64
1.2.2.4.5 Signal Preamplification . . . 66
1.2.2.4.6 Charge Injection Compensation . . . 67
1.2.2.4.7 Signal-to-noise Ratio . . . 70
1.2.2.4.8 Low-noise Operational Transconductance Amplifier (OTA) . . . 72
1.2.2.5 Decimation Filter . . . 74
1.2.2.5.1 Simple Counter . . . 74
1.2.2.5.2 Improved Decimation Filter . . . 75

1.2.3 Calorimetric Sensor . . . 75
1.2.3.1 Sensing Principle . . . 75
1.2.3.2 Calorimetric Transducer . . . 77
1.2.3.2.1 Thermally Insulated n-Well Island . . . 77
1.2.3.2.2 Aluminum/Polysilicon Thermopile . . . 77
1.2.3.3 Sensitivity of the Calorimetric Sensor . . . 78
1.2.3.4 Read-out Circuitry . . . 80
1.2.3.5 Low-noise Chopper Amplifier . . . 81
1.2.3.5.1 Introduction . . . 81
1.2.3.5.2 Signal Transfer Function . . . 82
1.2.3.5.3 Chopper and Demodulator . . . 83
1.2.3.5.4 Low-noise Preamplifier . . . 83
1.2.3.5.5 Bandpass Filter and Oscillator . . . 88
1.2.3.5.6 Third Amplifier . . . 93
1.2.3.5.7 Measurement Results . . . 93
1.2.3.6 Antialiasing Filter and A/D Converter . . . 96
1.2.3.7 Decimation Filter . . . 97
1.2.3.8 On-chip Calibration . . . 99

1.2.4 Resonant-beam Gas Sensor . . . 100
1.2.4.1 Sensing Principle . . . 100
1.2.4.2 Thermomechanical Actuation and Piezoresistive Detection . . . 101
1.2.4.2.1 Feedback Operation . . . 101
1.2.4.2.2 Cantilever Design . . . 102
1.2.4.2.3 Detection of the Cantilever Deflection . . . 102
1.2.4.2.4 Thermomechanical Actuation . . . 104

1.2.4.2.5 Thermal and Capacitive Crosstalk 106
1.2.4.3 Resonant-beam Oscillator 107
1.2.4.3.1 Feedback Topology 107
1.2.4.3.2 Differential Difference Amplifier (DDA) 108
1.2.4.3.3 Delay Line 112
1.2.4.3.4 High-pass Filter 113
1.2.4.3.5 Measurement Results 113
1.2.5 Acknowledgments 115
1.2.6 References 116

In this chapter, the use of industrial complementary metal oxide semiconductor (CMOS) technology (CMOS is the dominant technology for fabrication of microprocessors, memory, ...) with the aim of reducing costs, size and time to market of array-based gas detection systems is discussed. The first part was published in a previous volume of this series [1], and included current research on CMOS-based gas sensors and the design, fabrication, and system aspects of a single-chip gas detection system. Figure 1.2.1 shows a micrograph of this multisensor chip, which comprises three transducers that respond to fundamentally different analyte properties. The three sensors rely on commercially available polymeric layers to detect airborne volatile organic compounds (VOCs). This chapter, Part II, is focused on aspects concerning the single sensors. Section 1.2.1 gives an introduction to polymers as sensitive layers for gas detection. In subsequent sections, the three transducers and the corresponding on-chip circuitry are detailed.

1.2.1 Polymers as Sensitive Layers

The three transducers incorporated on the single-chip gas detection system rely on polymeric coatings to detect analytes in the gas phase. They record the changes in polymer properties upon absorption of the analyte. Figure 1.2.2 shows the working principle of the single-chip gas detection system. The transducers monitor changes in three different physical properties of the sensitive layer and convert these changes into an electrical signal, which is then processed by the electronics. The following section gives a short introduction to the transduction mechanism from the chemical domain into the physical domain. A detailed description of the thermodynamics of polymer-based chemical sensing can be found in [2] and [3], and physical chemistry aspects are detailed in [4].

1.2.1.1 Polymers for the Detection of VOCs

Volatile organic compounds (VOCs) usually comprise organic liquids with a boiling point below 250 °C (excluding methane). They include, eg, alcohols, aromatic compounds, and halogenated compounds, which are widely used in indus-

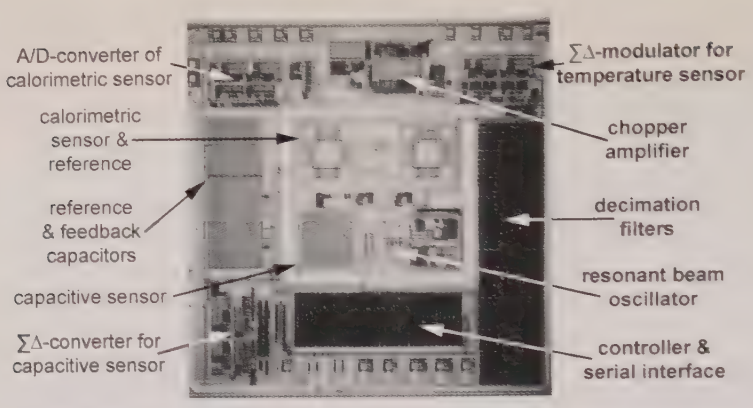


Figure 1.2.1. Micrograph of the single-chip gas sensor microsystem.

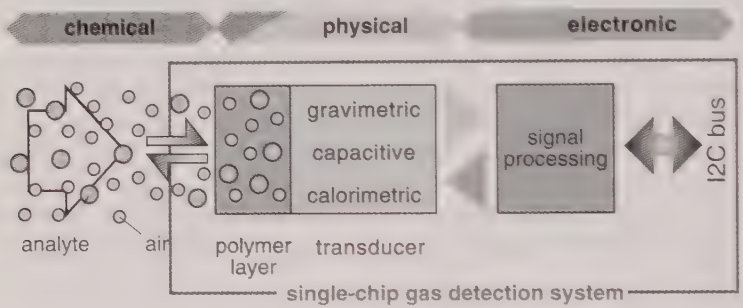


Figure 1.2.2. Working principle of polymer-coated microsensors for gas detection.

try. There is a strong demand for sensors detecting VOCs at ppm concentrations because VOCs are flammable (eg, dimethyl ether), carcinogenic (eg, benzene), or detrimental to the ozone layer (eg, halogenated compounds). Polymers are widely used as sensitive layers for the detection of VOCs for several reasons:

- A wide range of polymers can be bought off the shelf. These polymers have been developed for other than gas-sensing applications but their properties are well known, and they have proven to be long-term stable.
- Polymers can be easily deposited and structured using either spray coating with shadow masks or drop coating methods.
- The absorption and desorption of VOCs in polymers is in general completely reversible.

1.2.1.2 Polymer–Analyte Interaction

The single-chip gas detection system has two mechanisms that infer selective detection of analytes in gas mixtures. The first is the transduction principle that re-

sponds to a specific property of the analyte, eg, its mass. The second is the selectivity introduced by the specific polymer–analyte interaction.

Polymer–analyte interactions range from weak physisorption to strong chemical interactions such as charge transfer or covalent bonds (chemical reaction, 120–800 kJ/mol). Chemical reactions are selective but mostly not reversible and are, therefore, rarely used for gas sensors. As a rule of thumb, interactions with energies below 20 kJ/mol are mostly reversible at room temperature.

Standard polymers with proven long-term stability are used as sensitive layers. They rely on pure physisorption. All three transducers measure bulk properties of the polymer. Hence, for the thick polymer layers used here (> 1 μm), only ab-/desorption of the analyte into the bulk of the polymer has to be considered and surface adsorption of analyte molecules can be neglected.

The polymers used for the work described here are poly(etherurethane) (PEUT), poly(dimethylsiloxane) (PDMS), poly(methyloctylsiloxane) (PMOS), and poly[methyl(cyanopropyl)siloxane] (PMCPS). PEUT was purchased from Thermedics (Woburn, MA, USA) (SG-80A) and was used as the standard polymer to characterize and compare the different chemical sensors. Details on PEUT can be found in [5]. The other polymers are polysiloxane derivatives. They differ in the specific side chains connected to the Si–O group. Details on polysiloxanes can be found in [3]. The functionalized side chains influence the sensitivity of the polymer towards certain groups of analytes. Such 'selectivity patterns' are based on specific polymer–analyte interactions:

- *Dispersion:* London forces between any type of molecules. They account for the major part of all polymer–analyte interactions.
- *Dipole–dipole (dipole-induced dipole):* Between polar molecules. This can be used for partially selective detection of polar analytes in polar or polarizable polymers.
- *Hydrogen bonds:* Between a hydrogen atom covalently bound to atom A and a second atom B that must be an electron donor. The interaction energy is about 20 kJ/mol.

1.2.1.3 Modeling of the Polymer–Analyte Interaction

The absorption of an analyte into a polymer can be described by a single characteristic parameter, the partition coefficient K_c. K_c is defined as the ratio between the equilibrium concentration of the analyte in the polymer phase and its concentration in the gas phase:

$$K_c = \frac{c_{poly}}{c_{gas}} \tag{1.2.1}$$

The partition coefficient can be derived from the Gibbs free energy as a special case of the general equilibrium constant K. The change of the Gibbs energy G due to a chemical reaction [4] can be written as

$$\Delta G = \sum_i n_i \mu_i^0 + RT \sum_i n_i \ln a_i = \Delta G^0 + RT \ln \left(\prod_i a_i^{n_i} \right) \tag{1.2.2}$$

where μ_i is the chemical potential of component i and n_i is its stoichiometric number (or amount of substance), ΔG^0 is the standard reaction Gibbs energy and a_i is the activity. The activity a is also called the effective concentration and is related to the real concentration c via the equation $c_i = \gamma \cdot a_i$, where γ denotes the activity coefficient. In the case of physisorption, no chemical reaction between the analyte and the absorption sites in the polymer occurs and therefore the activity (or concentration) of absorption sites remains constant. Equation (1.2.2) can be simplified accordingly (the stoichiometric number n_i is positive for the products and negative for the reactants; see [4]). The change of the Gibbs energy G at equilibrium is zero and Equation (1.2.2) can be rewritten as

$$\begin{aligned} 0 &= \Delta G^0 + RT \ln \left(\frac{a_{\text{analyte}}^{\text{sorbed}}}{a_{\text{analyte}}^{\text{free}}} \right) = \Delta G^0 + RT \ln K_c \\ \ln K_c &= \frac{-\Delta G^0}{RT} \end{aligned} \tag{1.2.3}$$

According to the fundamental Gibbs equation, ΔG^0 is composed of an enthalpy term ΔH^0 and an entropy term ΔS^0. Physisorption of analytes into polymers can be conceptually modeled as a two-step process. First the analyte condenses into the liquid phase on the surface of the polymer and then the liquid analyte mixes with the polymer.

Equation (1.2.3) can then be written in the form

$$\ln K = \frac{\Delta H_{\text{vap}} - \Delta H_{\text{mix}} - T(\Delta S_{\text{vap}} - \Delta S_{\text{mix}})}{RT} \tag{1.2.4}$$

where the vaporization enthalpy ΔH_{vap} is the negative condensation enthalpy. The same consideration applies to the entropy of vaporization ΔS_{vap}. The enthalpy and entropy of vaporization (ΔH_{vap} and ΔS_{vap}) are in most cases the dominating terms, but they are not specific for the polymer–analyte combination. The partial selectivity results from the smaller mixing terms in Equation (1.2.4). Typical values for K_c are in the range 100–10000.

Within a certain class of analytes (homologous series), the mixing properties are similar. Hence some general conclusions for K can be derived from Equation (1.2.4) [3]:

- K_c decreases approximately exponentially with increasing operation temperature.
- At a defined temperature, K_c is inversely proportional to the analyte saturation vapor pressure and proportional to its boiling point.

1.2.2 Capacitive Sensor

1.2.2.1 Sensing Principle

The dielectric properties of an analyte can be assessed by monitoring the change in the polymer dielectric constant upon analyte absorption from the gas phase. The straightforward way to measure the dielectric constant is to measure the capacitance of a polymer-filled plate capacitor. The first publications on miniaturized polymer-based capacitive chemical sensors date back from the late 1960s. They describe sensors to measure relative humidity [6, 7] and (a few years later) VOCs [8]. The most popular approach to realize sensitive capacitors are polymer-coated interdigitated electrodes [7–9], because they are ideally suited for fabrication in a planar technology such as CMOS. Furthermore, planar polymer-coated capacitors provide direct access of the analyte to the sensitive layer. More complex sensor structures have been developed to, eg, reduce the response time [10].

A schematic diagram of the interdigitated capacitor is shown in Figure 1.2.3. The capacitors are fabricated exclusively with layers and processing steps available in the standard CMOS process sequence. Electrode E1 is made from the first metal layer while electrode E2 is a stack of the first and the second metal layers. The pad-etch is used to remove the passivation on top of the sensing capacitor. The computed electrode configuration was chosen to enhance the sensitivity of the polymer-coated sensor by maximizing the polymer volume in the region with strong electric field [5].

Figure 1.2.4a shows a micrograph of the interdigitated capacitor. It consists of 128 electrode pairs and occupies an area of $824 \times 814\ \mu m^2$. The scanning electron microscope (SEM) picture in Figure 1.2.4b shows a cross section of an uncoated sensing capacitor. The width and spacing of single electrodes are 1.6 μm and the electrode periodicity is 3.2 μm.

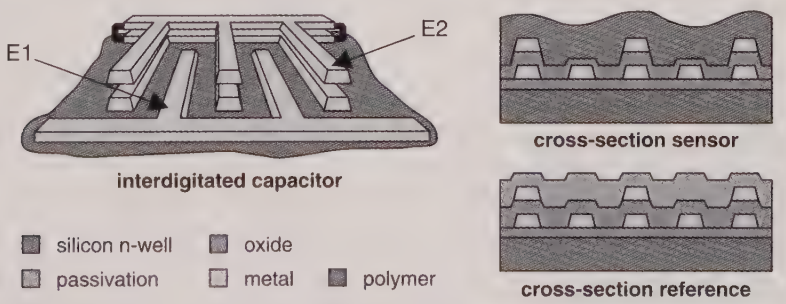


Figure 1.2.3. Schematic diagram of the sensing and reference capacitor made from the two metal layers of the CMOS process. Electrode E2 is a stack of Metal 1 and Metal 2. The sensitivity is increased by maximizing the polymer volume in the region with strong electric field.

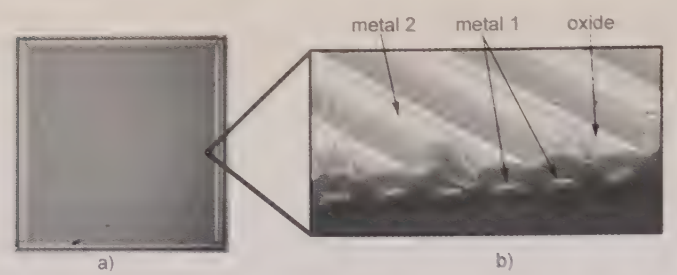


Figure 1.2.4. (**a**) Micrograph of $800 \times 800\ \mu m^2$ sensing capacitor with 128 interdigitated finger electrodes, the capacitor is split into two halves. (**b**) SEM of a cut through the sensing capacitor. The passivation on top of the metal 2 electrode has been removed using the pad etch of the CMOS process.

1.2.2.2 Equivalent Circuit Model

A schematic diagram of one electrode pair is shown in Figure 1.2.5a. In Figure 1.2.5b the equivalent circuit model for calculation and simulation is depicted. From finite element method (FEM) simulations of a sensing capacitor coated with 10 µm of PEUT (ε=2.97), the capacitances in the equivalent circuit model in Figure 1.2.5 have been determined [5]. The results are summarized in Table 1.2.1. For many polymers the conductivity can be neglected. The two capacitors C_{ox2} and $C_{polymer}$ can then be substituted by a single capacitor C_{poly}.

The changes in capacitance due to absorption of the analyte into the polymer are in the attofarad range. For a capacitor coated with 3.2 µm of PEUT, the sensitivity to toluene is ≈ 4 aF/ppm. In view of the small capacitance changes, on-chip signal conditioning circuitry is imperative [11].

Two effects have to be considered in order to describe the change in the sensing capacitor upon absorption of the analyte into the polymer. The first effect is the change of the dielectric constant $\varepsilon_{polymer}$ due to a difference between $\varepsilon_{analyte}$ and $\varepsilon_{polymer}$. The second includes the volume change of the polymer. The latter

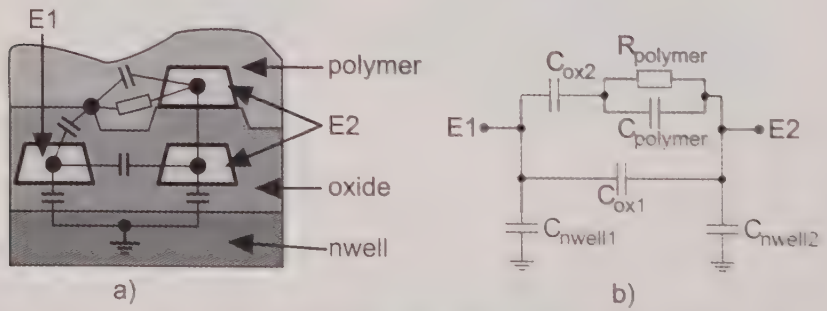


Figure 1.2.5. (**a**) Cross section of one electrode pair; (**b**) equivalent circuit model.

Table 1.2.1. Simulated values (pF) for the capacitors in Figure 1.2.4. C_{ox2} and $C_{polymer}$ were calculated from C_{poly} with the assumption that their ratios are equal to the ratios of the layer thickness for the oxide and the polymer

C_{poly}	C_{ox1}	C_{nwell1}	C_{nwell2}	C_{ox2}	$C_{polymer}$
1.4	6.4	17.7	18.2	–8.4	–1 .7

effect is called swelling and can be neglected if the polymer layer is thicker than the electrode pitch. A detailed analysis of the two effects can be found in [5]. In the following, only thick polymer layers will be considered.

The change of the capacitance C_{poly} upon absorption of analyte can then be described by

$$C_{poly} = C_{poly0}\left[1 + c_{gas} VF_C \left(\frac{\varepsilon_a}{\varepsilon_{polymer}} - 1\right)\right] \tag{1.2.5}$$

The dielectric constants of the analyte and polymer are denoted by ε_a and $\varepsilon_{polymer}$. C_{gas} is the concentration of the analyte in the gas phase. The relative volume fraction VF_C is defined as

$$VF_C = \frac{V_{analyte}}{V_{analyte} + V_{polymer}} \frac{1}{c_{gas}} \approx \frac{V_{analyte}}{V_{polymer}} \frac{1}{c_{gas}} \tag{1.2.6}$$

The relative volume fraction is proportional to the partition coefficient K introduced in Section 1.2.2.3:

$$VF_C = \frac{V_A}{n_{poly}} \cdot K \tag{1.2.7}$$

1.2.2.3 Read-out Circuitry

Micromachined capacitive sensors have been developed for many different applications, eg, accelerometers, pressure sensors, fingerprint sensors, and gas sensors. Various read-out circuitry topologies have been developed to measure the small capacitors of micromachined sensors. Most of the designs record the difference between a sensing capacitor and a reference capacitor not affected by the measurand. As a consequence, these designs do not provide accurate information about the absolute value of the capacitor. This offers the advantage that some parasitic effects such as temperature drift, and ageing affect the reference capacitor to the same extent as the sensor and hence cancel out.

The topologies used for fingerprint sensors [12] do not satisfy the resolution requirements for gas sensing applications. Some implementations for accelerometers use sine-wave or square-wave excitation of capacitive half-bridges in

Table 1.2.2. Comparison of different topologies for capacitive microsensor interfaces

Author	Application	Resolution	Minimum detectable capacitance	Bandwidth	Sensitive to parasitics
ADXL202	Accel.	10	50 aF	10	No
Jachowicz	Humidity	10	20 fF	N	Yes
Lemkin	Accel.	13	20 aF	100	No
Kung	General	12	50 aF	260	No
Malcovati	General	15	50 aF	50	No

combination with analog demodulation techniques to measure the capacitance [13–15]. Sensors based on displacement (eg, pressure sensors, accelerometers) of one of the capacitor plates can use force feedback to compensate for the change in capacitance [16]. This method circumvents sensor nonlinearities but cannot be applied to polymer-coated interdigitated capacitors because no feedback mechanism is available. Relaxation oscillators using the sensing capacitor as frequency-determining element have also been developed [17]. The most popular approach to read-out capacitive sensors with high resolution and good suppression of parasitics is switched capacitor design [18–22]. Table 1.2.2 compares the performances of the different designs.

A resolution of 19 bits is needed in order to achieve a detection limit of 1 ppm for VOCs. The bandwidth can be as low as 1 Hz. As none of the listed designs exhibits such performance, an improved architecture based on a switched capacitor $\Sigma\Delta$ modulator was developed.

1.2.2.4 Second-order Sigma–Delta Modulator

1.2.2.4.1 Single-bit $\Sigma\Delta$-Modulation Techniques for Sensor Interfaces

Single-bit $\Sigma\Delta$ modulators have become very popular for A/D conversion in sensor interfaces [23]. Their main advantage over other converters is that no high-precision elements dividing the reference voltage are needed. The large oversampling ratio required for high accuracy $\Sigma\Delta$ converters is not a limiting factor because the bandwidth of many physical and chemical sensors is low. The basic architecture of $\Sigma\Delta$ modulators (see Figure 1.2.6) can be adapted in many ways. This allows for improved topologies where the sensors are directly incorporated into the modulator in order to increase the performance or decrease the complexity of the interface circuitry. One approach is to replace the loop filter by the sensor element because many active sensors have a first- or second-order low-pass transfer function. This technique was first employed in capacitive accelerometer designs with force feedback [16, 22] and later adapted for thermal flow sensors [24] and microfluxgate sensors [25].

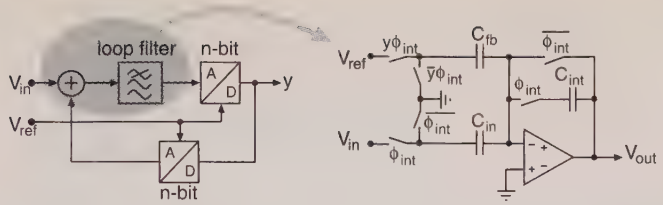


Figure 1.2.6. Modulator architecture for sensor applications. Basic architecture of a $\Sigma\Delta$ modulator (left) and switched-capacitor integrator used as loop filter for first-order $\Sigma\Delta$ modulator (right).

In the case of the capacitive sensor, no filtering function or feedback mechanism is available. Most $\Sigma\Delta$ modulators use a switched-capacitor integrator as shown in Figure 1.2.6 to realize the first stage of the loop filter. The charge added to the integration capacitor C_{int} in each clock-phase is $C_{in} \times V_{in}$. Wherease in the case of a common A/D converter the input voltage changes and C_{in} are constant, a capacitive sensor can be realized when C_{in} is replaced with the sensing capacitor and V_{in} is kept constant. This method was used by Malcovati [22]. In this section, improvements and changes to the basic architecture are discussed.

1.2.2.4.2 Second-order Sigma–Delta Modulator

A system that measures the difference between a sensing and a reference capacitor (C_S and C_R) can be realized by converting the integrator in Figure 1.2.6 (right) into a fully differential design where the two input capacitors $C_{in1,2}$ are replaced by sensor and reference. Fully differential design is preferable owing to the better power-supply rejection ratio and the first-order cancellation of charge injection. A straightforward implementation of such an integrator is shown in Figure 1.2.7.

The problem with the circuit in Figure 1.2.7 is its charge-transfer efficiency: While the common-mode voltage at the output is controlled by the common-mode feedback of the fully differential amplifier, the input common-mode voltage VCM_{in} is defined by the capacitive divider between the sensing capacitor C_S

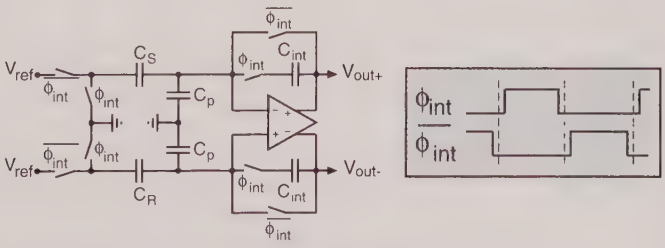


Figure 1.2.7. Straightforward implementation of a switched-capacitor integrator to convert the difference between C_S and C_R.

and the other capacitors. If we assume $C_S = C_R$ and an initial voltage of zero on both integration capacitors, $VCM_{in}[n]$ is given by

$$VCM_{in}[n] = V_{ref}\frac{C_S}{C_S + C_{int} + C_p}\sum_{k=0}^{n}\left(\frac{C_{int}}{C_S + C_{int} + C_p}\right)^k$$
$$VCM_{in}[\infty] = V_{ref}\frac{C_S}{C_S + C_p} \tag{1.2.8}$$

The change in the output voltage of the integrator in each clock-cycle is given by

$$\Delta V_{out} = \frac{(V_{ref} - VCM_{in})(C_S - C_R)}{C_{int}} \tag{1.2.9}$$

If no parasitic capacitance C_p is present at the input of the operational transconductance amplifier (OTA), the resulting sensitivity is zero. In the case of the capacitive chemical sensor with large parasitics and large input transistors for the OTA, 75% of the maximum sensitivity is obtained.

Therefore, a different topology with VCM_{in} equal to VCM must be used to achieve the full sensitivity and avoid contributions of the parasitic capacitors. There are two modes of charge transport in a switched-capacitor amplifier: In the inverting mode the capacitor is discharged during the phase ϕ_{int} and connected to V_{ref} during the integration phase ϕ_{int}. The noninverting mode uses a clocking scheme where the two phases are exchanged. This can be used to solve the problem with the common-mode input voltage without an additional voltage source.

In the first integrator of the modulator shown in Figure 1.2.8, the sensing and the reference capacitor were split in two parts. One half is operated in the invert-

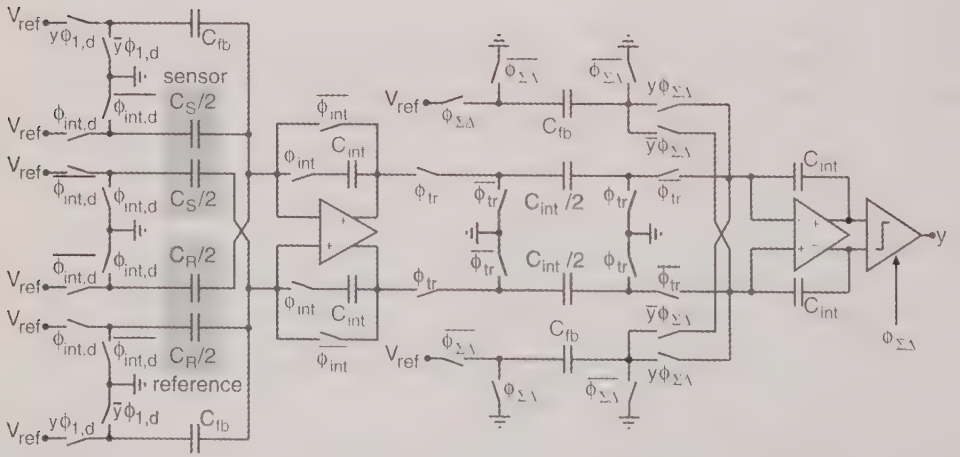


Figure 1.2.8. Schematic diagram of the second-order switched-capacitor ΣΔ modulator.

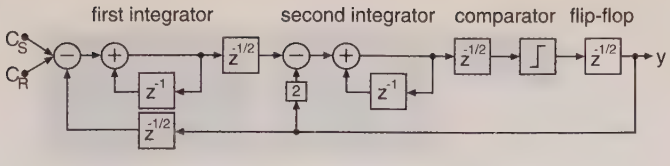


Figure 1.2.9. Signal-flow graph of the ΣΔ modulator.

ing mode and connected to the negative input of the OTA while the other half operates in the noninverting mode and is connected to the positive input. In this way, the common-mode voltage is always zero, and the charge that is transferred to the integration capacitor is $C_S \times V_{ref}$. The resulting voltage change at the output of the OTA after one clock cycle is then given by

$$\Delta V_{out} = \frac{V_{ref}(C_S - C_R)}{C_{int}} \tag{1.2.10}$$

Correlated double sampling (CDS) is used to eliminate the input offset of the OTA and to reduce the $1/f$ noise. The offset of the OTA is sampled onto the input capacitors in the CDS-phase ϕ_{int}. This voltage is inverted and transferred to the integration capacitor C_{int} during the integration phase ϕ_{int} and cancels the new offset contribution. Compared with other topologies that have been developed for CDS [21], this method does not need any additional capacitors to store the input offset. This reduces the area overhead of the CDS and eliminates matching errors, since the charge is always stored on the same capacitor.

Figure 1.2.8 shows a schematic diagram of the complete second-order ΣΔ modulator. The second integrator is a standard switched-capacitor integrator because no correlated double sampling is needed. The signal flow graph of the ΣΔ modulator is shown in Figure 1.2.9. The output for an input signal $X = C_S - C_R$ in the z-domain is then given by

$$Y = z^{-3/2}(X + Q(1 - z^{-1})^2) \tag{1.2.11}$$

where Q is the quantization error added by the comparator. The signal-transfer function is unity. The desired second-order noise shaping of the quantization error is obtained.

1.2.2.4.3 Tunable Reference Capacitor

An ideal reference capacitor is exactly identical with the sensing capacitor and has no sensitivity towards the analyte. This can only be achieved when sensor and reference are both coated with the same polymer and the reference is sealed from the gas flow. Unfortunately, an appropriate packaging solution for this method is not available. Therefore, only the sensor is coated with polymer while

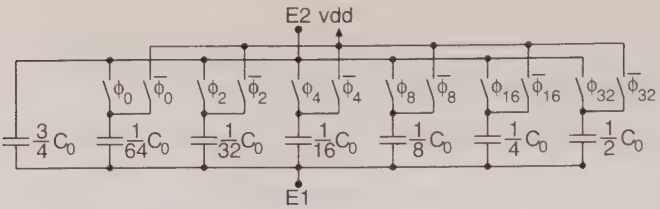

E2 vdd
ϕ_0 $\bar{\phi}_0$ ϕ_2 $\bar{\phi}_2$ ϕ_4 $\bar{\phi}_4$ ϕ_8 $\bar{\phi}_8$ ϕ_{16} $\bar{\phi}_{16}$ ϕ_{32} $\bar{\phi}_{32}$
$\frac{3}{4}C_0$ $\frac{1}{64}C_0$ $\frac{1}{32}C_0$ $\frac{1}{16}C_0$ $\frac{1}{8}C_0$ $\frac{1}{4}C_0$ $\frac{1}{2}C_0$
E1

Figure 1.2.10. Schematic diagram of the tunable reference capacitor.

the reference is shielded from the gas flow by the passivation layer (silicon nitride) of the CMOS process. For polymers with high dielectric constants, a large difference between sensing and reference capacitor results. In the worst case, this offset exceeds the dynamic range of the $\Sigma\Delta$ modulator.

The design in [22] uses two external adjustable voltage sources connected to the input of the sensing and the reference capacitor to compensate this offset ($V_{\text{sensor}} \times C_S = V_{\text{reference}} \times C_R$). A third adjustable voltage source V_{ref}, which provides the bias for the feedback capacitors, is used to adjust the dynamic range of the $\Sigma\Delta$ modulator. As the noise from these voltage sources is not correlated, it degrades the overall noise performance of the $\Sigma\Delta$ modulator. Drift of one of the references translates into errors of the output signal. Furthermore, three digitally adjustable on-chip voltage sources occupy a lot of area and increase the power consumption.

Therefore, a single reference voltage is used to bias the sensing, reference, and feedback capacitors (see Figure 1.2.8). The noise contributions in the three signal paths are thus correlated and cancel out. The effects of reference voltage drift are also diminished. If V_{sensor} equals $V_{\text{reference}}$, a different solution has to be developed, which compensates for the initial difference between sensing and reference capacitor. The reference capacitor is split into a fixed and a tunable part. The tunable part consists of six binary weighted capacitors, which have been assembled from 63 interdigitated unit capacitors (see Figure 1.2.10). The tuning range of C_0+50%/–25% is sufficient to calibrate the sensor for all the polymers used as a sensitive layer.

1.2.2.4.4 Interdigitated Feedback Capacitors

The bitstream from the $\Sigma\Delta$ modulator is proportional to the ratio of input and feedback capacitors (C_{fb}). Poly/poly feedback capacitors were used in [22]. The temperature coefficient (TC) of the interdigitated capacitors does not match the TC of the poly/poly capacitors, which results in a large temperature sensitivity. Interdigitated capacitors are used for C_{fb} to match the temperature drift (see Figure 1.2.11). This also simplifies the calibration of the system because the ratio of two interdigitated capacitors is well defined, while the ratio of an interdigitated capacitor and a poly/poly capacitor is subject to the large process spread of CMOS processes.

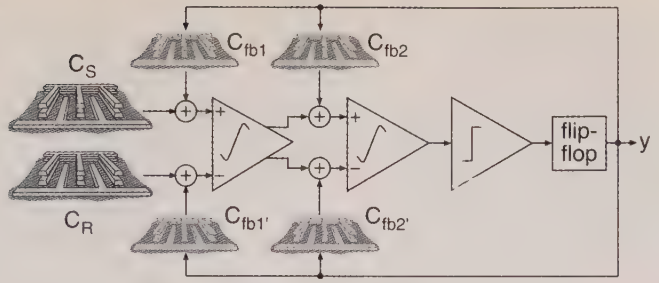


Figure 1.2.11. Block diagram of the $\Sigma\Delta$ capacitance-to-digital converter. The sensor, reference, and feedback capacitors have been realized as interdigitated capacitors in order to reduce the temperature drift of the converter.

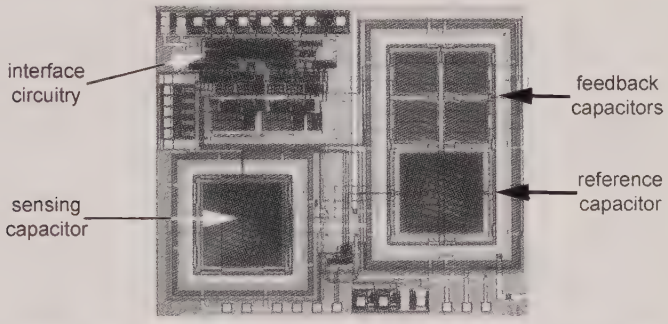


Figure 1.2.12. Test chip with interdigitated sensing, reference, and feedback capacitors. All capacitors are located on thermally insulated n-well islands to allow measurements at different sensor temperatures [5].

The nonlinearity of the interdigitated capacitors is large compared with poly/poly capacitors. This does not affect the performance of the one-bit $\Sigma\Delta$ modulator, since all interdigitated capacitors have the same bias voltage V_{ref}. The integration capacitors C_{int} must be linear because their voltages change in every clock cycle. All integration capacitors are therefore realized as poly/poly capacitors.

To characterize the various temperature effects, a chip with interdigitated sensing, reference, and feedback capacitors was designed [26]. Figure 1.2.12 shows a micrograph of an etched and uncoated chip. The passivation was removed on all interdigitated capacitors to allow for identical coatings on the sensing and reference capacitors. Furthermore, all interdigitated capacitors were located on thermally insulated n-well islands. A polysilicon heating resistor was integrated on the membranes to control the temperature of the n-well island. This allows for measurements at different sensor temperatures [5]. The chip also includes a multiplexer, which enables to select either poly/poly or interdigitated feedback capacitors. The test chip was used to compare the temperature performance of a design with poly/poly feedback capacitors with a design with interdigitated feed-

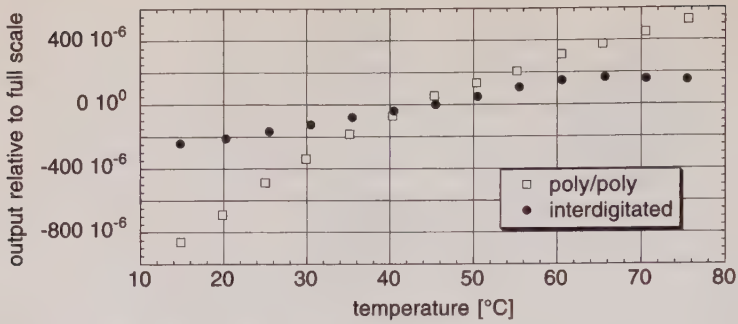


Figure 1.2.13. Relative temperature drift of the switched-capacitor $\Sigma\Delta$ modulator with poly/poly feedback capacitors and interdigitated feedback capacitors.

back capacitors. All interdigitated capacitors were covered with Katiobond 4670 (Delo, Germany) to shield them from external factors such as humidity. The results are shown in Figure 1.2.13. For the interdigitated feedback capacitors a drift of 5 ppm/°C was measured. This is three times less than for the poly/poly feedback capacitors.

1.2.2.4.5 Signal Preamplification

The changes of the sensing capacitances are in the attofarad range and small compared with the absolute value of the interdigitated capacitors. Therefore, only a small fraction of the input range of the A/D converter is used. This leads to complex and area-consuming designs of the decimation filters. The input range of the modulator is defined by the feedback capacitors and their biasing voltage. The biasing voltage is equal for all interdigitated capacitors (see Section 1.2.2.4.2) and cannot be modified to reduce the input range. A reduction in the dynamic range can be achieved by designing the feedback capacitors four times smaller than the sensing capacitor. Smaller sizes are not appropriate because the noise of the modulator would increase and the matching of the capacitors would degrade.

A different approach has been chosen to increase the input signal without adding any additional switches or amplifiers in the signal path. By operating the integrator in the first stage of the $\Sigma\Delta$ modulator at a higher clock frequency than the rest of the converter, the signal can be amplified. Figure 1.2.14 shows the different clock signals for an amplification factor of four. With every cycle of the integrator clock ϕ_{int}, the charge difference $V_{\text{ref}} \times (C_S - C_R)$ is accumulated on the integration capacitor C_2. If the conversion takes place at a 2^n times lower clock rate $\phi_{\Sigma\Delta}$, the output signal is given by

$$Out = 2^n \frac{C_S - C_R}{C_{fb}} \tag{1.2.12}$$

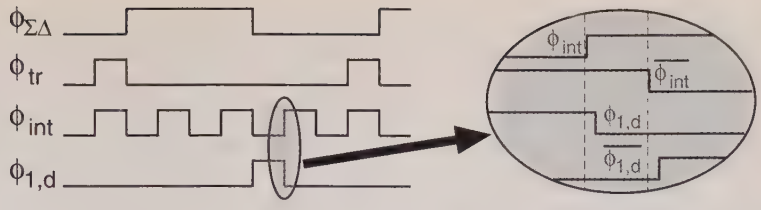


Figure 1.2.14. Clocking scheme for the $\Sigma\Delta$ modulator with a preamplification factor of 4 ($n=2$).

Amplification factors from 1 to 8 can be selected in the current design by setting the values of n from 0 to 3. The additional digital cells necessary to generate the clock signals consume only about 3% of the total chip area.

This topology also minimizes the power consumption of the sensor system. Only the first integrator operates at high clock frequency. The second integrator, the comparator, and the decimation filter operate at a 2^n times lower clock frequency.

1.2.2.4.6 Charge Injection Compensation

When a transistor is turned off in a switched capacitor circuit, the charge stored in the channel and the parasitic capacitors is distributed to the capacitors on each end of the switch (see Figure 1.2.15). This effect is called charge injection and leads to errors in the output voltage of the integrator. This problem was already identified in the first switched-capacitor circuits [27] and has been analyzed in detail in later publications [18, 28–30].

The charge stored in the channel of a minimum size switch (0.8 µm CMOS process, $V_{dd}=5$ V) leads to error voltages of a few millivolts on a 1 pF capacitor. These errors can be minimized by optimizing the modulator itself and the design of the switches. The following paragraphs describe the compensation techniques that have been used for the design of the capacitive sensor.

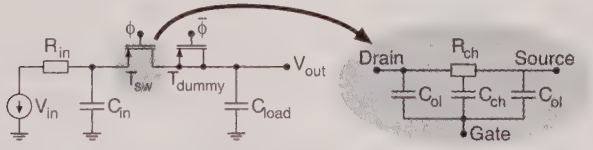


Figure 1.2.15. Charge injection in analog MOS switches. The transistor can be modeled by a channel resistance R_{ch} and a distributed channel capacitance C_{ch} [31]. C_{ol} denotes the overlap capacitance for drain and source, respectively.

Modulator Topology

- *Fully differential design:* In a fully differential design the charge, which is injected into the positive branch of the integrator, is identical with the charge injected into the negative branch. Therefore, charge injection can be seen as a common-mode signal, which is suppressed by the common-mode feedback. Unfortunately, the matching of the charge injected by the switching transistors is bad (typically, mismatches are in the range of 10%) and the charge injection must be further reduced.

- *Signal-dependent charge injection:* Charge injection depends on the input voltage of the switch. This leads to different amounts of charge injected into the two branches of the fully differential integrator. In a standard $\Sigma\Delta$ modulator with voltage input this cannot be avoided. In the case of a capacitance input, all input voltages can be made identical. This eliminates the effect of signal-dependent charge injection (see Figure 1.2.8). A similar technique is used for the switch connected to the integration capacitor C_{int}. If the switch is connected to the output of the OTA, the input voltage of the switch changes in every clock-cycle and signal-dependent charge injection is observed. When the switch is connected to the input of the OTA, the input voltage of the switch is always VCM.

- *Delayed switching:* The timing of the switches can also be used to reduce the effects of charge injection [31]. The switching of the input switches is delayed with respect to the switch connected to the integration capacitor C_{int} at the end of the integration phase ϕ_{int}. Therefore, only the charge injection of the switch connected to C_{int} has to be taken into account, while the charge from the input switches only affects the parasitic capacitors at the input of the OTA. The same procedure can be applied to reduce charge injection on the input capacitors C_{S} and C_{R} during the phase ϕ_{int}, but this is not implemented in the current design.

Design of the Switches

- *Dummy switch compensation:* The most popular way to compensate charge injection is to connect a half-size transistor in a metal oxide semiconductor (MOS) capacitor configuration on the drain side of the switch (see Figure 1.2.15) and operate this switch on the inverted switching clock [27]. Assuming that 50% of the charge stored in the switching transistor goes to the load capacitor C_{load}, the charge injection from the switching transistor is exactly cancelled.

- *Slope of switching transient:* The redistribution of the charge stored in the channel of the two switching transistors strongly depends on the slope of the switching transient [28]. For long transient times the charges are distributed according to the capacitor ratios on either side of the switch in Figure 1.2.15. For short transients the charges are equally divided on each side regardless of the capacitive loads. The latter case must be achieved, if the compensation

with half size dummy transistors should be effective. From the equations given in [28], a maximum switching time of 2 ns was calculated for an error of less than 5% from the ideal charge distribution and assuming that $0.3 < C_1/C_2 < 3$.

- *Minimum size switches:* The charge stored in the channel of a transistor is directly proportional to its gate area. Therefore, minimum size transistors are used to reduce charge injection. The minimum size of the switching transistors is restricted by the maximum on-resistance that can be tolerated in order to achieve full settling of the integrator in one clock period. In the current design, NMOS transistors with a size of 4/2 µm are needed to fulfill this requirement.
- *n-Type (NMOS), p-type (PMOS), and complementary metal oxide semiconductor (CMOS) switches:* NMOS transistors are the best choice for switching transistors, because they have the lowest on-resistance with respect to a given gate area. For nodes at voltages larger than $(V_{dd}-V_{th})$ or for nodes with large voltage swing, PMOS switches or combined CMOS switches are needed. In the current design only the input switches (to allow for different reference voltages) and the reset switch of the modulator (to ensure reliable startup) are realized as CMOS switches. For CMOS switches, a different switching scheme had to be developed in order to achieve minimum charge injection over the complete input voltage range. The best results were obtained for a sequential clocking scheme, where the NMOS switch and its dummy transistor are switched off first (delay 1 ns). After a settling time of 5 ns, the PMOS transistor and its dummy transistor are switched off. The remaining error voltage due to charge injection for different input voltages is shown in Figure 1.2.16.

Several different methods to reduce further the effects of charge injection have been proposed but they are not suitable for the design of a switched-capacitor $\Sigma\Delta$ modulator.

- An error amplifier has been developed [29] that measures the charge injection and adapts the input voltage of the dummy transistor in order to minimize the charge injection. This method needs one operation amplifier and two large capacitors for each switch and is too costly for applications where many transistors need to be compensated.

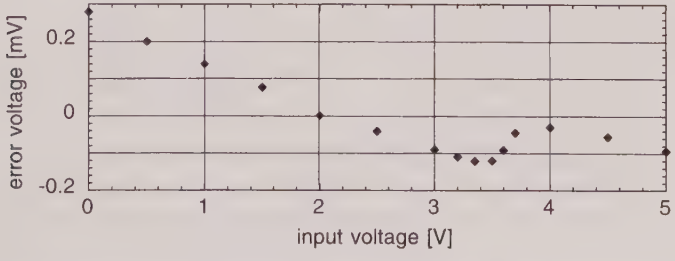


Figure 1.2.16. Error voltage for different input voltages in a CMOS switch with dummy-transistor compensation (simulation).

- A technique for designs that need large switches to satisfy the settling requirements has been presented [12]. A parallel combination of a large and a small transistor is used. The large switching transistor is turned off first. The charge injection from this transistor has no effect, because the small transistor is still on. After an appropriate delay, the small transistor is turned off and the remaining charge injection is compensated with a half-size dummy transistor. In the $\Sigma\Delta$ modulator for the capacitive sensor this method would lead to only small improvements, because the switching transistors are already close to the minimum design rules.

1.2.2.4.7 Signal-to-noise Ratio

The noise in a modulator has two sources: quantization noise and electronic noise from switches and OTAs.

The quantization noise is generated in the comparator that is used as a one-bit A/D converter. The noise shaping in the feedback of the modulator has been derived [32]. The noise power transfer function (NTF) of a second-order modulator is equal to the second term in Equation (1.2.11).

$$NTF[z] = |(1 - z^{-1})^2|^2$$
$$NTF[\omega] = 6\left(1 - \frac{4}{3}\cos(\omega T) + \frac{1}{3}\cos(2\omega T)\right) \tag{1.2.13}$$

The total input-referred electronic noise is dominated by the contribution from the first integrator. The noise of the second integrator and the comparator is divided by the gain of the first integrator, which is large in the signal band. These sources can therefore be neglected. The noise of the first integrator has been calculated using the circuit in Figure 1.2.17, where all the relevant noise sources are drawn.

The largest contribution comes from the input-referred noise of the operational amplifier (V_{op}). The noise of the operational amplifier has to be analyzed sepa-

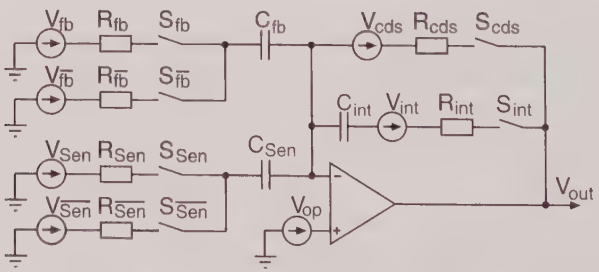


Figure 1.2.17. Equivalent circuit model of the first integrator with all relevant noise sources. The fully differential integrator can be replaced with a single-ended design because it is completely symmetric.

rately because it is affected by the correlated double sampling scheme (CDS), which is used to cancel the input offset of the amplifier [29, 33]. The model shown in Figure 1.2.17 is based on the model described in [34]. The modified version described here uses two different transfer functions for the noise sampled throughout the reset phase and the integration phase to account for the different switching configurations. The sampled noise power on the integration capacitor C_{int} can then be derived from the model in Figure 1.2.17

$$H_{\text{CDS}}[\omega] = H_{\text{Int}}[\omega] - H_{\text{Reset}}[\omega] \cdot \mathrm{e}^{-j\omega(T_s/2)}$$

$$S_{\text{CDS}}[\omega] = \tau^2 \left(\frac{\sin\left(\pi\tau \dfrac{\omega}{\omega_s}\right)}{\pi\tau \dfrac{\omega}{\omega_s}} \right)^2 \sum_{k=-\infty}^{\infty} \left(S_{\text{OpAmp}}[\omega - k \cdot \omega_s] \cdot |H_{\text{CDS}}[\omega - k \cdot \omega_s]|^2 \right) \tag{1.2.14}$$

where H_{Int} and H_{Reset} are the transfer functions from the input of the operational amplifier to C_{int} during the respective clock phases, is the hold time of the sampled signals, T_s is the clock-period, and S_{OpAmp} is the equivalent input noise power density (double sided) of the operational amplifier:

$$S_{\text{OpAmp}}[\omega] = \frac{1}{2} 4kT R_{\text{Op}} \left(1 + \left| \frac{\omega_{\text{corner}}}{\omega} \right| \right) \tag{1.2.15}$$

R_{Op} is a calculated equivalent resistor for the input referred thermal noise floor of the operational amplifier. R_{Op} equals 3 kΩ, which corresponds to a noise level of 7 nV/Hz$^{1/2}$. The corner frequency of the amplifier is 30 kHz.

$$f_{\text{corner}} = \omega_{\text{corner}}/(2\pi) \tag{1.2.16}$$

The noise of the switches is not correlated. Therefore, the input-referred noise contributions from the single switches were derived separately for each switch. As only the 'on' phase of the switch is relevant for the noise calculation, each switch has to be considered only once, either during the integration phase or the CDS phase. The contribution from the integration phase is obtained by first calculating the voltage sampled on the integration capacitor C_{int} and then dividing the result by the transfer function. For the CDS phase, the voltage on C_{int} is calculated from the voltage sampled on the input-capacitors C_{Sen} and C_{fb} and is then divided by the transfer function.

The calculated noise contributions of the noise sources in Figure 1.2.17 are shown in Figure 1.2.18. The noise of the operational amplifier accounts for approximately 80% of the total noise power, while the other 20% are shared between the switches. This uneven split is due to the large parasitic capacitor of the sensor at the input of the operational amplifier. The voltage on these parasitic capacitors is modulated only by the noise source of the operational amplifier, which causes this noise to be amplified by the respective capacitance ratio.

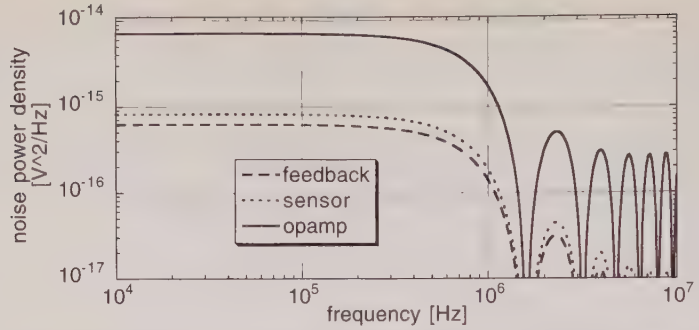


Figure 1.2.18. Calculated noise contributions from the single noise sources in the first integrator of a second-order $\Sigma\Delta$ modulator.

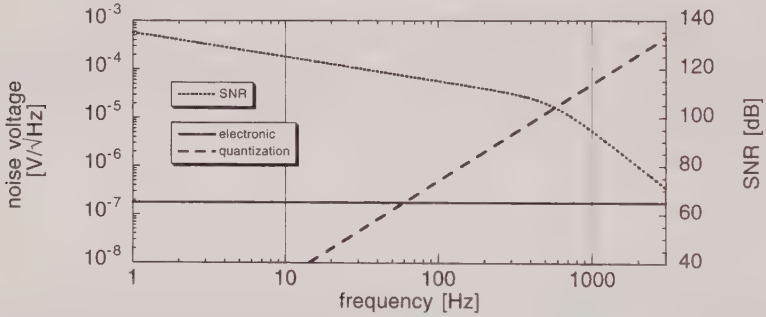


Figure 1.2.19. Noise and SNR of the second-order $\Sigma\Delta$ modulator. For frequencies up to 70 Hz, the electronic noise dominates the overall noise. The SNR increases by 40 dB/decade down to a bandwidth of 500 Hz.

The total noise of the $\Sigma\Delta$ modulator is shown in Figure 1.2.19. The electronic noise dominates for frequencies up to 70 Hz. The resulting signal-to-noise ratio (SNR) for the respective bandwidth is also shown in Figure 1.2.19. When electronic noise dominates, the SNR increases only by 3 dB for each doubling of the oversampling ratio. This is the same amount that is achieved, when the bandwidth of Nyquist-rate A/D converters is reduced. For frequencies with the quantization noise as largest contribution, the second-order noise shaping of the $\Sigma\Delta$ modulator becomes effective, and the SNR increases by 12 dB for each doubling of the oversampling ratio.

1.2.2.4.8 Low-noise Operational Transconductance Amplifier (OTA)

Specifications

From the considerations above, some requirements for the OTA used in the $\Sigma\Delta$ modulator can be derived. While the multisensor chip uses a sampling frequency

Table 1.2.3. Specifications of the operational transconductance amplifier

AV0	>90 dB	Integrator leakage
Gain-bandwidth product (GBW)	>30 MHz	Sampling frequency
Thermal noise floor	$<10\ nV/Hz^{1/2}$	SNR
1/*f*-noise corner-frequency	30 kHz	SNR
Slew rate	>5 V/μs	Settling behavior
Output swing	>±3 V	Input range

of 800 kHz, the ΣΔ modulator was designed for operation up to 2 MHz in order to enable measurements of the dielectric properties at higher frequencies. Measurements at different sampling frequencies are desirable, because some features of the analyte can be derived from the spectral response of the dielectric constant. The most important specifications of the ΣΔ modulator have been summarized in Table 1.2.3.

Design

Figure 1.2.20 shows a schematic diagram of the OTA. A fully differential folded-cascode amplifier with a PMOS input stage was chosen, because this topology can meet the gain requirements without adding a second amplification stage. PMOS transistors were chosen for the input stage, because of the lower 1/*f* noise. The inputs of the common-mode feedback (CMFB) are the NMOS-biasing transistors T_{2x}. The gain–bandwidth product (GBW) of the common-mode feedback must be larger than the GBW of the differential amplifier [35]. The main difficulty in this design included meeting this GBW requirement without degrading the noise performance of the amplifier. The bandwidth requirement leads to the condition $gm_2 > gm_1$ [35] for the standard fully differential folded-cascode OTA.

The noise contribution of transistor *T*2 relative to the input transistor *T*1 is given by

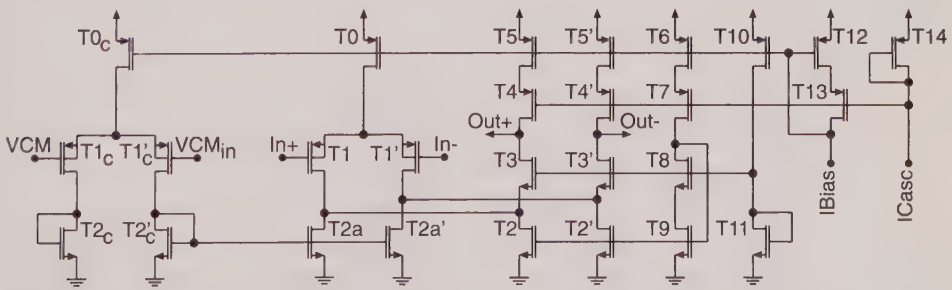


Figure 1.2.20. Schematic diagram of the operational transconductance amplifier.

Table 1.2.4. Transistor sizes and bias currents used for the OTA

Transistor	W (μm)/L (μm)	Current source	Current (μA)
$T1$	600/2	$T0$	2×200
$T1c$	300/2	$T0c$	2×50
$T2$	150/10	$T5$	100
$T2a$	20/3	$T12$	50
$T2c$	40/3		
$T5$	300/10		

$$\begin{aligned} \frac{P_{\text{th}}[T2]}{P_{\text{th}}[T1]} &= \frac{gm_2}{gm_1} && \text{thermal noise} \\ \frac{P_{1/f}[T2]}{P_{1/f}[T1]} &= \frac{KF_{\text{NMOS}}}{KF_{\text{PMOS}}}\left(\frac{L_{T1}}{L_{T2}}\right)^2 && \text{flicker noise} \end{aligned} \tag{1.2.17}$$

where $L_{1,2}$ is the length of the respective transistor and $KF_{\text{NMOS, PMOS}}$ is the flicker noise constant given by the technology. The contribution of transistor $T2$ is minimized by decreasing gm_2 and increasing L_2. On the other hand, gm_2 must be as large and L_2 as small as possible in order to meet the bandwidth requirements. These competing requirements are the key trade-off in designing the operational amplifier.

In the design shown in Figure 1.2.20, the noise of the differential amplifier and the GBW of the common-mode amplifier can be set independently by splitting the NMOS transistor $T2$ into two parts and adding a differential gain stage at the common-mode input of the amplifier. The GBW of the CMFB is then given by

$$GBW_{\text{CM}} = \frac{gm_{1\text{fb}}}{2\pi C_{\text{L,CM}}}\,\frac{gm_{2\text{a}}}{gm_{2\text{C}}} \tag{1.2.18}$$

If transistor $T1$ dominates the thermal and flicker noise, the GBW can be optimized independently of the noise. The transistors sizes and the necessary bias currents are summarized in Table 1.2.4. The performance of the OTA was verified by simulation and measurement of a separate test chip. The experimental results for both are in good agreement with the calculated values.

1.2.2.5 Decimation Filter

1.2.2.5.1 Simple Counter

Equilibrium-type measurements have to be conducted with the capacitive sensor integrated on the single-chip gas sensor system, because the transient behavior is not used for the discrimination. Therefore, a bandwidth of 1 Hz is sufficient. An integrated decimation filter with a resolution of 19 bits occupies a large area

($4\ mm^2$ is occupied by the 13-bit decimation filter for the calorimetric sensor). A simple counter was hence used to decimate the bitstream coming from the $\Sigma\Delta$ modulator. The transfer function of the counter can be described by a moving-average filter [36] followed by a downsampler. The sampling rate is reduced by the ratio of the decimation counter gate time to the sampling period. The SNR of the converter for a sampling frequency of 800 kHz and a gate time of 1.3 s is 112 dB, which corresponds to a resolution of better than 18 bits.

1.2.2.5.2 Improved Decimation Filter

It is important for some applications to look at transient signals while equilibrium conditions are established. In order to achieve a resolution of 19 bits for signals up to 50 Hz, a better decimation filter is needed. Throughout this work, the bitstream from the test chips was acquired with a fast digital PC-card and then filtered with standard digital filters available in different software packages (eg, LabView™, Matlab™, etc.).

1.2.3 Calorimetric Sensor

1.2.3.1 Sensing Principle

A definition of a calorimetric gas sensor can be found in [2]: 'Calorimetric sensors rely on determining the presence or concentration of a chemical by measurement of an enthalpy change produced by the chemical to be detected. Any chemical reaction or physisorption process releases or absorbs a certain quantity of heat from its surroundings'. For sensors coated with standard polymers, only physisorption contributes to the enthalpy change, because no chemical bonds between polymer and detected analyte are formed. The calorimetric sensor principle is shown in Figure 1.2.21. When the analyte is absorbed into the polymer,

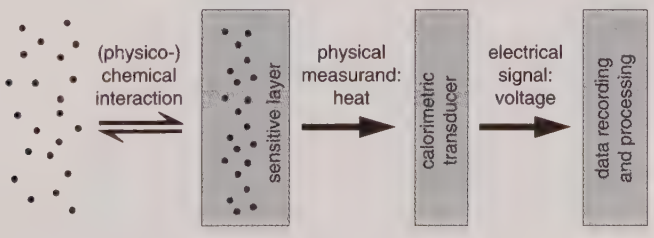


Figure 1.2.21. Principle of the calorimetric gas sensor.

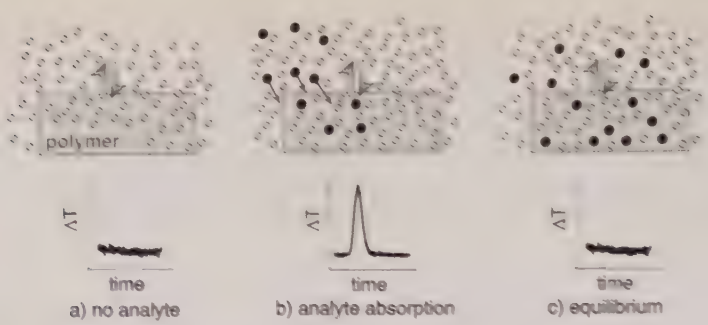


Figure 1.2.22. Transient signal response of the calorimetric sensor.

the heat of condensation and the heat of mixing lead to a temperature increase, ΔT, in the polymer:

$$\Delta T \sim \Delta H: \quad \Delta H = \Delta H_{\text{cond}} + \Delta H_{\text{mix}} \tag{1.2.19}$$

ΔH is the overall enthalpy change (in J/mol). The first term on the right-hand side of Equation (1.2.19) is the enthalpy change due to analyte condensation. The second term describes the enthalpy change due to the mixing of analyte and polymer. For many VOC–polymer combinations the mixing term is small in comparison with the overall enthalpy change and can sometimes be neglected [3, 29].

At thermodynamic equilibrium, the change in the Gibbs free energy $\Delta G = \Delta H - T \cdot \Delta S$ is zero (ΔS denotes the change in entropy). Therefore, only a transient signal is obtained from the calorimetric sensor. This is illustrated in Figure 1.2.22. The initial state is thermal equilibrium and no analyte present in the air, hence no temperature signal is shown in Figure 1.2.22a. Analyte in ambient air is then absorbed by the polymer and the condensation and mixing enthalpy temporarily increase the temperature of the polymer (Figure 1.2.22b). As soon as an equilibrium concentration is established, the signal returns to zero, because there is no net enthalpy change (Figure 1.2.22c, dynamic equilibrium: equal number of molecules absorb and desorb). This has some implications for the use of thermopiles in gas-sensing applications:

- Slow changes in the ambient analyte concentration are difficult to detect, because they lead to a continuous, but extremely small, change of enthalpy. Therefore, a switching scheme has to be implemented, where the analyte-loaded air is alternated with a reference gas (eg, synthetic air) in order to generate fast concentration changes.
- In a switched system, the amplitude of the signal depends largely on the diffusion processes in the gas manifold and the diffusion of the analyte into the polymer. Therefore, the peak area of the integral signal has to be used to obtain information about the total enthalpy change upon concentration changes.

1.2.3.2 Calorimetric Transducer

1.2.3.2.1 Thermally Insulated n-Well Island

The temperature changes in the polymer are in the millikelvin range. Therefore, the measurement area has to be thermally insulated from the silicon substrate, which is an excellent thermal conductor. Membranes made from the dielectric layers of the CMOS process have been used for infrared sensor arrays. Dielectric membranes show a parabolic temperature profile. A flat temperature profile over the sensitive area is, however, desired for gas-sensing applications. This can be achieved by creating a thermally insulated n-well island as shown in Figure 1.2.23. The island is then coated with the sensitive polymer.

1.2.3.2.2 Aluminum/Polysilicon Thermopile

The temperature on the membrane can be assessed with different transducers using, eg, the temperature coefficient of bipolar transistors or resistors. The sensitivity of these approaches is not suitable to measure temperature changes in the millikelvin range. A differential measurement between membrane and substrate is preferred over an absolute measurement, because it eliminates the influence of ambient temperature fluctuations. This can be done using thermocouples, which are based on the Seebeck effect. The hot contacts are located on a thermally insulated n-well island while the cold contacts are placed on the substrate. The best detectivity for thermocouples is obtained using bismuth and antimony [37]. These materials are not available in CMOS processes and would be difficult to deposit and pattern. From the conductors available in a CMOS process, the poly-

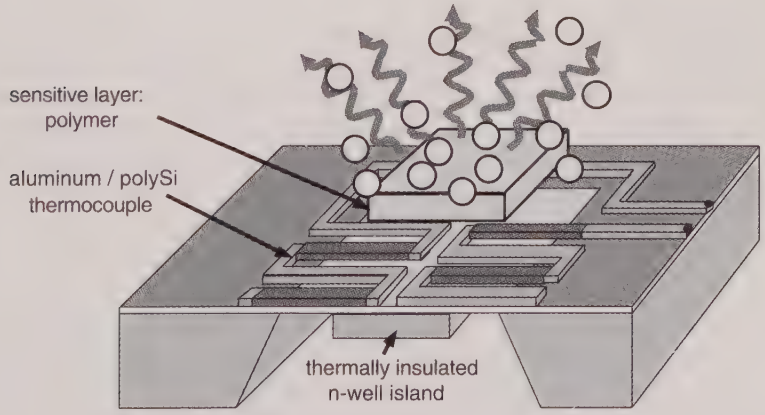


Figure 1.2.23. Polymer-coated thermally insulated n-well island.

silicon/aluminum thermocouple exhibits the best sensitivity of 110 μV/K. In order to measure temperature differences in the millikelvin range, many thermocouples must be connected in series for enhanced sensitivity.

1.2.3.3 Sensitivity of the Calorimetric Sensor

The sensor system has to be optimized for a maximum SNR at the output of the A/D converter while allowing for measurements of a wide concentration range of different analytes. The trade-offs that are relevant for the design of the overall system are outlined in this section. The implications for the read-out circuitry and the chosen design will be described in detail. The transducer optimization is beyond the scope of this discussion. For calorimetric gas sensors, results can be found in [38, 39]. Detailed studies on the closely related design of thermal imagers based on suspended thermocouples have been performed [37, 40].

The overall sensitivity of the calorimetric sensor (for thin polymer layers) is described by

$$S_{\text{calorimetric}} = \frac{\Delta V_{\text{thermopile}}}{\mathrm{d}c_{\text{gas}}/\mathrm{d}t} = A \cdot \Delta\gamma \cdot N \cdot K_{\text{c}} \cdot \Delta H \cdot V_{\text{polymer}} \tag{1.2.20}$$

where ΔH is the enthalpy change per mole given by Equation (1.2.19), V_{polymer} is the volume of the polymer, $\Delta\gamma$ is the difference of the Seebeck coefficient of the thermopile materials (see below) and K_{c} is the partition coefficient. The constant A (K s/J) includes the feature sizes of the transducer and material properties. The sensitivity can be split into two parts: the chemical and the physical sensitivity.

Chemical Sensitivity

The chemical sensitivity is defined as the heat generated by an analyte concentration change in the gas phase. It is determined by the analyte–polymer combination, which finally provides the selectivity needed to discriminate different analytes. The contributions to the chemical sensitivity are the partition coefficient K, the enthalpy change ΔH, and V_{polymer}. More information on the chemical sensitivity can be found in [3].

Physical Sensitivity

The physical sensitivity is the output voltage of the system in response to a defined heating power on the membrane. It includes the features of the transducer but does not depend on the analyte. The first transduction mechanism involved is the temperature increase due to the applied power. The parameters to be optimized are as follows:

- *Size of the membrane:* A large membrane leads to a large sensitive area and a large number of thermocouples. The mechanical stability of the membranes and the maximum allowable chip-area determine the size limits. The membranes have sizes of $650 \times 650\ \mu m^2$ (square membrane, see Figure 1.2.1) and $2150 \times 750\ \mu m^2$ (rectangular membrane on the chip developed for glob-top packaging [1]), respectively.
- *Spacing between substrate and n-well island:* Larger spacing improves the thermal isolation of the n-well island and, therefore, the efficiency of the heating. The mechanical stability of the membrane and the maximum chip area limit this spacing. The increase in the thermocouple resistance due to the larger distance between hot and cold contact also reduces the improvement.
- *Number and width of thermocouples:* The signal increases linearly with the number of thermocouples. However, the thermocouples consist of aluminum and polysilicon, both of which exhibit a high thermal conductance. This reduces the thermal resistance of the membrane and therefore degrades the heating efficiency. These considerations suggest the use of minimum-width conductors to increase the thermal resistance and make efficient use of the available area. On the other hand, wide conductors are needed in order to minimize the thermal noise caused by the low electrical conductivity of the polysilicon resistors. The square membrane has a total of 132 thermocouples, the rectangular membrane features 300.

The thermopile converts the temperature difference between substrate and n-well island into a voltage signal. For the best overall performance, the input-referred SNR of the first amplifier, not the signal, has to be maximized. This input-referred SNR of the first amplifier is described by

$$\mathrm{SNR} = 20 \cdot \log \frac{N \cdot \Delta\gamma \cdot \Delta T}{\sqrt{4 \cdot k \cdot T \cdot (R_{\mathrm{th}} + R_{\mathrm{eq}}) \cdot B}} \tag{1.2.21}$$

where N is the number of thermocouples, $\Delta\gamma$ is the difference of the Seebeck coefficients of the two materials forming the thermocouples, ΔT is the temperature difference, R_{th} is the resistance of the thermopile, R_{eq} is the equivalent input noise resistance of the read-out circuitry, and B is the signal bandwidth. The terms in the numerator of Equation (1.2.21) have already been considered above. In the following paragraphs, the noise contributions in the denominator of Equation (1.2.21) will be analyzed. The two thermopiles used here have resistances of 40 kΩ (square membrane) and 143 kΩ (rectangular membrane).

For high source resistance, the optimization of the circuitry can be decoupled from the optimization of the sensor, if the contribution of the circuitry is negligible. In order to reduce the additional noise voltage inferred from the amplifier to below 10%, the equivalent input noise resistance needs to be below 8.4 kΩ (20 kΩ for the rectangular membrane). This corresponds to a white noise of $12\ \mathrm{nV/Hz^{1/2}}$ ($18\ \mathrm{nV/Hz^{1/2}}$). For designs with small source resistance, this

approach cannot be chosen. The overall system optimization includes the input noise of the first amplifier as well as the power and area consumed by the complete read-out circuitry [37].

The second important parameter is the signal bandwidth, which must be as small as possible in order to minimize the noise. From experimental data one can conclude that the highest frequency of interest in the transient signal during absorption or desorption of analytes is about 400 Hz.

1.2.3.4 Read-out Circuitry

From the aforementioned considerations, the requirements for the read-out circuitry are derived. The resulting system shown in Figure 1.2.24 includes a differential arrangement, where a polymer-coated sensor is connected in series with an uncoated reference. Temperature differences caused by flow and radiation are cancelled by the differential arrangement. The small signals are then amplified by a low-noise chopper amplifier. Chopping was chosen because of the better noise performance compared with correlated double sampling (CDS). The chopping frequency is 5 kHz. The gain of the amplifier must be adjustable in order to allow measurements over a large concentration range and the use of different analyte polymer combinations. The maximum gain of 6400 can be reduced by a factor of 4 or 16. Together with a resolution of 12 bits this is sufficient for the target applications. An antialiasing filter succeeding the amplifier is needed to avoid undersampling of high frequency noise by the A/D converter.

The system has a narrow signal band from 0 to 400 Hz. The requirements for the stop band are set by noise constraints, by the suppression of the signals upconverted to the chopping frequency by the demodulator at the output of the amplifier, and by clock feedthrough. A minimum damping of 80 dB at the chopping frequency is needed to meet these requirements. By using a largely oversampled $\Sigma\Delta$ A/D converter, the filtering can be shifted to the digital decimation filter, where low frequencies do not increase complexity, area, and power consumption of the design. This also simplifies the design of the antialiasing filter. Owing to the narrow signal band, an oversampling factor of 128 can easily be realized. This makes a one-bit $\Sigma\Delta$ modulator the most suitable choice for the 12-bit A/D converter owing to the less stringent matching requirements.

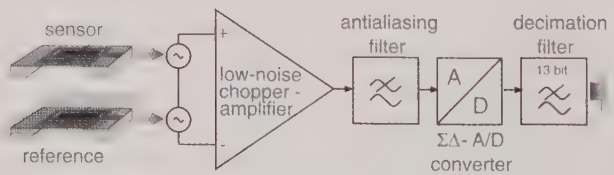


Figure 1.2.24. Block diagram of the read-out circuitry of the calorimetric sensor.

Table 1.2.5. Important design parameters for the read-out circuitry

Component	Parameter	Value
Amplifier	Chopping frequency	5 kHz
	Bandwidth	500 Hz
	Adjustable gain	6400/1600/400
	Equivalent input noise	<12 nV/Hz$^{1/2}$
	Residual offset	<10 V
Antialiasing filter	Damping at $f_S/2$	>40 dB
	Gain	0 dB
2nd-order $\Sigma\Delta$ A/D modulator	Sampling frequency	100 kHz
	SNR at bandwidth 400 Hz	74 dB
Digital decimation filter	Accuracy	12 bit
	Bandwidth	400 Hz
	Output sampling rate	800 kHz/ 1024 = 781 Hz
	Damping at f_{chopp}	>40 dB

The most important specifications for the different building blocks of the read-out circuitry are summarized in Table 1.2.5.

1.2.3.5 Low-noise Chopper Amplifier

1.2.3.5.1 Introduction

Low offset and low noise are difficult to achieve for designs in standard CMOS technology owing to several problems:

- Large mismatch in CMOS input stages: The equivalent input offset of CMOS input stages cannot be reduced below 1 mV even with common-centroid layout techniques.
- Large $1/f$ noise in the input stage of amplifiers.
- Accurate and large resistors are not available in a standard CMOS process.

In [41], several techniques to reduce offset and noise of an amplifier are compared. The two most popular methods are correlated double sampling (CDS) and chopping. For low-noise, low-bandwidth instrumentation amplifiers, the chopper amplifier is the best choice because of its superior noise performance [40, 42]. The architecture of the chopper amplifier is determined by the following system requirements:

- In order to achieve a gain of 6400 without trimming and an overall accuracy of better than ±5%, three gain stages are needed because it is difficult to realize ratios larger than 20 with such performance.

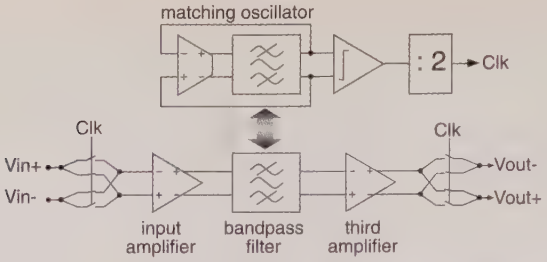


Figure 1.2.25. Block diagram of a low-noise chopper amplifier with a gain of 6400.

- The amplifier has a gain of up to 6400. Owing to the large gain, the offset of the first amplifier (~ 1 mV) cannot be simply upconverted by the demodulator, because it would saturate the amplifier. A bandpass filter after the first amplifier is used to eliminate this offset. This bandpass filter also reduces the residual offset caused by the charge injection spikes generated in the input chopper [42].
- The nested-chopper technique proposed in [41] uses a second chopper signal to reduce further the residual offset to 100 nV. The trade-off is a reduced bandwidth (~ 5 Hz). This is not sufficient for the calorimetric sensor, which requires a bandwidth of 400 Hz. A residual offset of 10 μV can be tolerated for this application, because the input signal shows an offset of the same order owing to incomplete cancellation of flow-effects.

After considering all the issues listed above, the architecture as shown in Figure 1.2.25 was selected [42]. Fully differential design was chosen, because it reduces the influence of charge injection and crosstalk from the noisy digital part of the chip. The overall gain was distributed equally between the two amplification stages and the filter. The gain of the first amplifier and the filter can be reduced by a factor of 4 using a digital selection signal. In the following section the different building blocks of the amplifier will be described in more detail.

1.2.3.5.2 Signal Transfer Function

The chopper amplifier shown in Figure 1.2.25 modulates the input signal to the chopping frequency and its harmonics using a simple cross-coupled switch. The chopped signal is first amplified and then bandpass filtered. The bandpass filter is used to remove the DC offset of the first amplifier, to reduce the residual offset [42], and to provide additional gain. The third amplifier is needed to achieve the final gain of 6400. The bandpass filter not only removes higher order harmonics of the charge injection spikes caused by the chopper amplifier but also filters the harmonics of the signal. The overall gain of the amplifier after demodulation is thus given by

$$Av_{\text{chopped}} = \frac{8}{\pi^2} \cdot Av_{\text{preamp}}[\omega 0] \cdot Av_{\text{bandpass}}[\omega 0] \cdot Av_{\text{thirdamp}}[\omega 0] \qquad (1.2.22)$$

where $Av_{preamp}[\omega 0]$, $Av_{bandpass}[\omega 0]$, and $Av_{thirdamp}[\omega 0]$ denote the gains of the respective amplifiers at the chopping frequency $\omega 0/(2\pi)$.

1.2.3.5.3 Chopper and Demodulator

The choppers at the input and output of the amplifier are realized with cross-coupled switches as shown in Figure 1.2.25. The size of the switches is determined by a trade-off between on-resistance (decreases with increasing *W/L* ratio) and charge injection. As the amount of charge is directly proportional to the size, minimum size transistors would be preferable in a single-ended configuration. In a fully differential configuration, however, the difference in charge injection between the two branches is determined by the matching of the two transistors. This suggests a slightly larger transistor size. A size of 20/1 [μm/μm] was chosen to achieve an on-resistance that is smaller than 10 kΩ under all biasing conditions and for all corners of the process parameters.

The chopping frequency [42] is determined by a trade-off between bandwidth and filtering requirements on the one side and the residual offset of the system on the other side. The residual offset is proportional to the chopping frequency. Filters are needed to remove the signal and noise components at the chopping frequency after demodulation and cut-off signals at frequencies exceeding the desired bandwidth. Hence the minimum chopping frequency is given by the signal bandwidth and the filter requirements. Leaving some margin for future increases in the bandwidth and accounting for the large process tolerances of CMOS processes, the chopping frequency was set to 5 kHz.

1.2.3.5.4 Low-noise Preamplifier

The most important constraints for the first amplifier are low noise, low power consumption, large input resistance, and a gain that is large enough to relax the noise specifications for the subsequent stages. A large bandwidth is needed to avoid phase shifts between the modulator and the demodulator. The amplifiers input offset must be smaller than 5 mV to avoid saturation of the bandpass filter. The input and output swing is not a major concern: For a minimum gain of 400 for the overall amplifier, the maximum differential input voltage for the preamplifier is 5 mV.

The use of operational amplifiers or differential difference amplifiers (DDAs) with feedback resistors is not feasible, because the large gain bandwidth and the small feedback resistors consume too much power. Therefore, an amplifier (without feedback) employing transconductance ratios of transistors to define the gain was designed. The first version of the amplifier is based on the design in [42] and uses transistors in strong inversion. An improved version further reduces power consumption by using transistors biased in the weak inversion region.

Strong Inversion

Figure 1.2.26 shows the schematic of the design using transistors in strong inversion. The gain is defined by the ratio of the transconductances of T_2, T_{2a} of the input stage and the linearized transconductance formed by transistors T_6 and T_7:

$$Av = \frac{gm_2}{gm_{67}} = \frac{gm_2}{gm_6/\left(1 + \frac{\beta 6}{4\beta 7}\right)} \tag{1.2.23}$$

The input stage of the preamplifier is split in two parts to allow the gain to be switched between 20 and 5. Transistors $T_{0,1,2}$ are three times smaller than transistors $T_{0a,1a,2a}$. Therefore, when switch $G5$ is closed, the transconductance gm_2 of the input stage is reduced by a factor of 4.

A folded configuration was chosen to allow for the use of PMOS transistors for both transconductances. This improves the matching and therefore the accuracy of the gain. The degeneration transistor T_7 is needed for two reasons. First, the output swing of the amplifier is 200 mV, which is beyond the linear range of the transconductor. Linearization can be achieved by degeneration [43]. The circuit in Figure 1.2.26 uses the diode-connected transistor T_6 in series with transistor T_7 and was analyzed in detail in [44]. For increasing input currents, transistor T_7 changes from the triode region to operation in the saturation region:

$$\begin{aligned}
&\frac{i}{I_{\text{bias}}} = \frac{v}{aVd_{\text{sat}}}\sqrt{1 - \left(\frac{v}{2aVd_{\text{sat}}}\right)^2}; \quad |v| < 1.21\ Vd_{\text{sat}} \\
&\frac{i}{I_{\text{bias}}} = \frac{\pm\left(\frac{v}{Vd_{\text{sat}}}\sqrt{4a-2} \pm \sqrt{4a - 1 - \left(\frac{v}{Vd_{\text{sat}}}\right)^2}\right)^2}{(4a-1)^2}; \\
&\qquad 1.21\ Vd_{\text{sat}} < |v| < 2.83\ Vd_{\text{sat}}
\end{aligned} \tag{1.2.24}$$

where Vd_{sat} is the saturation voltage of transistor T_6, the factor a equals $1+\beta 6/(4\times\beta 7)$ with $\beta_x = \mu_p \times C_{ox} \times W_x/L_x$ and I_{bias} is the current through transistor T_6. The values of $1.21\ Vd_{\text{sat}}$ for the transition from the linear region to saturation and $2.83 Vd_{\text{sat}}$ for saturation of the transconductance were calculated assuming that the factor a equals 2.5. For $a = 2.5$ the simulated nonlinearity of the transconductance is below 1% for currents up to 80% of the bias current.

Second, the parameters that can be varied to define the gain of the amplifier are the W/L ratio and the biasing currents. As gm follows the square root of both terms, their product must be 400 in order to achieve a gain of 20. The degeneration transistor T_7 reduces the combined transconductance (T_6 combined with T_7) by a factor of 2.5. The scaling factor of the transistor dimensions and the bias current is thus reduced to 64.

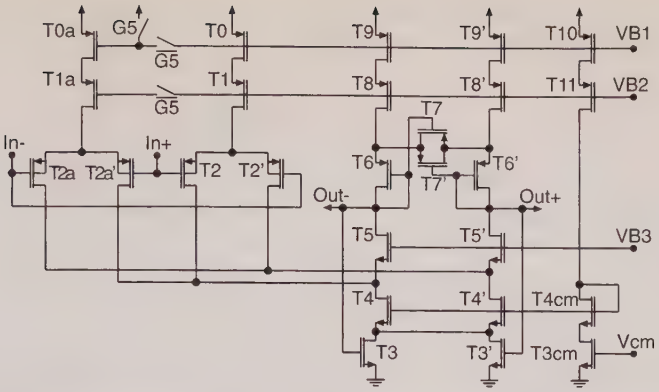


Figure 1.2.26. Schematic diagram of the low-noise preamplifier using transistors in strong inversion. The biasing voltages VB_{1-3} are generated by current mirrors. V_{CM} is the common-mode voltage.

Table 1.2.6. Summary of the design parameters of the transconductors used for the preamplifiers

Transistor	Strong inversion		Weak inversion		
	Size (μm/μm)	Current (μA)	Transistor	Size (μm/μm)	Current (μA)
T2	800/6	120	T2	1600/2	50
T3	60/15	150	T3	60/15	60
T4	66/25	150	T4	99/25	60
T6	60/8	30	T6	24/20	10
T7	10/8				

The bandwidth of the amplifier is given by the load transconductance gm_{67} and the capacitance at the output node:

$$f_{3\text{dB}} = \frac{1}{2\pi} \cdot \frac{gm_{67}}{C_{\text{par}} + C_{\text{L}}} = 2\ \text{MHz} \tag{1.2.25}$$

where C_{par} is the sum of the parasitic capacitances of the transistors at this node and C_{L} is the input capacitance of the subsequent filter. At a chopping frequency of 5 kHz this leads to a phase shift of only 0.5°. Capacitive compensation to achieve a lower bandwidth by pole splitting is not necessary, because the amplifier is operated under open-loop conditions. The sizes of the transistors and the respective bias currents are summarized in Table 1.2.6.

The noise performance is mostly determined by the input transistor *T2* and the bias transistor *T4*. The contribution of the transistors forming the load transcon-

ductance is less than 5% and can therefore be neglected. The input-referred thermal noise power density S_{th} is then given by

$$S_{th} = \frac{8}{3} kT \frac{1 + gm_4/gm_2}{gm_2} \tag{1.2.26}$$

The flicker noise is given by

$$S_{1/f} = \frac{1}{gm_2^2} \left(\frac{KF_p \cdot I_{T2}^{AF_p}}{f \cdot C_{ox} \cdot L_2^2} + \frac{KF_n \cdot I_{T4}^{AF_n}}{f \cdot C_{ox} \cdot L_4^2} \right) \tag{1.2.27}$$

where KF_p and KF_n are the flicker noise constants for the PMOS and NMOS transistors, respectively, AF is a parameter to fit the current dependence (provided by the foundry), C_{ox} denotes the gate capacitance per unit area, and L_x is the length of transistor T_x. This leads to a trade-off because reducing the noise by making gm_2 bigger increases either the power consumption or the area consumed by the input stage. On the other hand, while gm_4 should be small in order to reduce the noise contribution of transistor $T4$, a large gm_4 is needed for the CMFB (see below). PMOS transistors were chosen for the input stage, because they show better flicker-noise performance than NMOS transistors. At the chopping frequency, the thermal noise of transistor $T2$ contributes 36%, the flicker noise of transistor $T4$ 22%, the thermal noise of transistor $T4$ 13%, and the flicker noise of transistor $T2$ 4% to the total noise power. The thermal noise floor is at 7 nV/Hz$^{1/2}$ and the corner frequency is 3 kHz.

The common-mode feedback is based on the transistors T_3 and T'_3 operating in the linear region [45]. Their gates are driven by the output nodes $Out+$ and $Out-$. The sum of the currents is proportional to the CM voltage because the drains of T_3 and T'_3 are connected. Transistor T_4 serves as a cascode transistor for the CMFB and the current is then amplified by the active load formed by transistors T_8 and T_9. The problem with this type of CMFB is to achieve a GBW that is comparable to the bandwidth of the differential amplifier:

$$GBW_{CM} = \frac{1}{2\pi} \cdot \frac{gm_3}{C_{par} + C_{L,CM}} \cdot \frac{gm_4}{gds_3 + gm_4} \tag{1.2.28}$$

where C_{par} is the sum of the transistor capacitances at the output node (with respect to common mode signals) and $C_{L,CM}$ is the load capacitance needed to stabilize the CM amplifier. Stabilization is necessary because the CM amplifier is used with unity-gain feedback and the second pole caused by transistor T_4 has to be taken into account. The third term in Equation (1.2.28) is due to the current divider formed by the output conductance gds_3 of transistor T_3 and the transconductance of transistor T_4. As transistor T_3 is operating in the linear region, the output conductance is large and reduces the GBW. A large drain-source voltage Vds_3 increases gm_3 and decreases gds_3, but this voltage is limited by other design constraints. Vds_3 was set to 150 mV and the GBW is 3 MHz. Transistors

T_{4cm} and T_{3cm} are used to match the DC value of the common-mode output-voltage to the voltage VCM, which is externally applied to the gate of transistor T_{3cm}.

Weak Inversion

Transistors in weak inversion are the best choice for optimizing a design for noise performance versus low power and area consumption, because they offer the largest *gm/Id* ratio. If the basic topology of the circuit in Figure 1.2.26 is used, the load transistors must also operate in weak inversion in order to achieve good matching and an accurately defined gain over temperature and process variations. The transconductances of load transistors in weak inversion only allow a small voltage swing. The transition from weak inversion to strong inversion occurs at [35]

$$Vgs_{ws} = Vt_0 + 2\frac{kT}{q} \tag{1.2.29}$$

From Equation (1.2.29), a maximum voltage swing of 50 mV (at 300 K) per diode-connected load transistor results. The circuit in Figure 1.2.27 uses two stacked transistors in combination with a splitting technique [43] to achieve an output swing of ±400 mV, which is sufficient for the maximum input voltage of 5 mV.

The transconductance of transistors in weak inversion is directly proportional to the drain current. The gain of the amplifier is simply defined by a current ratio, which can be accurately set by matched current mirrors:

$$Av_{DC} = \frac{gm_2}{4gm_6} = \frac{Id_2}{4Id_6} = 20 \tag{1.2.30}$$

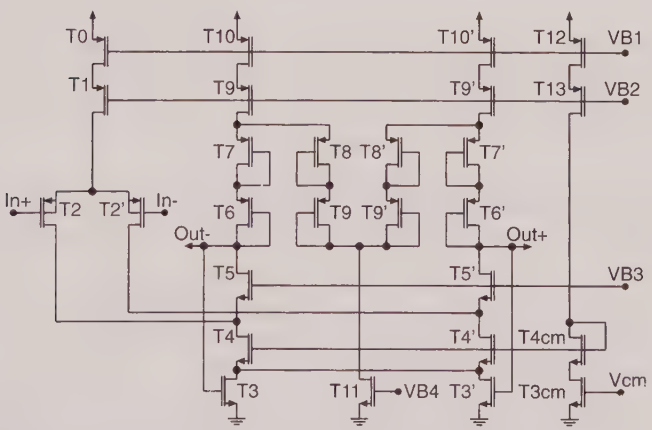


Figure 1.2.27. Schematic diagram of the preamplifier using transistors in weak inversion. The biasing voltages VB_{1-3} are generated by current mirrors.

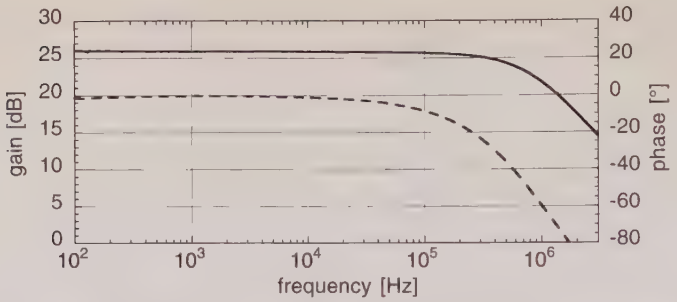


Figure 1.2.28. Measured transfer function of the weak inversion preamplifier (solid line, gain; dashed line, phase).

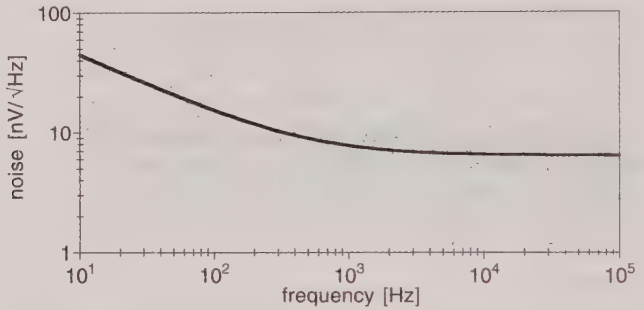


Figure 1.2.29. Measurement of the input-referred noise of the weak-inversion preamplifier.

The sizes and bias currents of the transistors are summarized in Table 1.2.6. The measured transfer function of the amplifier is shown in Figure 1.2.28.

The bandwidth is 800 kHz, the DC gain is 26.1 dB, and the phase shift at f_{chopp} is –0.8°. The variation of the gain with respect to bias-current variations over one order of magnitude (500 nA–5 μA) was measured to be 0.3 dB. The noise is again determined by the input pair and the NMOS bias transistor *T*4. Together these transistors account for 94% of the total simulated noise of 6.4 nV/$Hz^{1/2}$ (corner frequency 2 kHz). This is in good agreement with the measured value of 6.5 nV/$Hz^{1/2}$.

Compared with the strong inversion preamplifier, the weak inversion design achieves the same performance (gain, offset, noise) at a 2.5 times lower power consumption (900 μW versus 1.95 mW at V_{dd}=5 V).

1.2.3.5.5 Bandpass Filter and Oscillator

Bandpass Filter

Although the absolute value of the center frequency of the bandpass filter is not important, it must be accurately matched to the chopping frequency. Any mismatch between the two reduces the overall DC gain of the chopper amplifier by [42]

$$Av[\varepsilon] = \frac{Av[\varepsilon = 0]}{1 + (2Q\varepsilon)^2} \tag{1.2.31}$$

where Q denotes the quality factor of the bandpass filter, and is the relative mismatch between filter and chopping frequency:

$$\varepsilon = \frac{f_{\text{chopp}} - f_{\text{BP}}}{f_{\text{chopp}}} \tag{1.2.32}$$

The bandpass filter also reduces the residual offset of the amplifier by filtering the higher order harmonics in the wide-band spectrum of the spikes. Equation (1.2.33) compares the residual offset reduction of a design using a bandpass filter to a design employing a wide-band amplifier with infinite bandwidth:

$$k_{\text{Offset}} = \frac{\text{Offset}_{\text{BP}}}{\text{Offset}_{\text{Infinite}}} \approx \frac{8Q}{\pi} \frac{1}{1 + 8Q^2\varepsilon} \tag{1.2.33}$$

The exact equation for the offset reduction is derived in [42]. A detailed analysis of the effect of different filters on the residual offset can also be found there. This leads to an optimum mismatch of approximately –0.5% where the residual offset is theoretically zero (exact matching has been targeted in this design). Further filter requirements include a noise level below the output noise of the preamplifier and an adjustable gain of 5 or 20. These requirements can be met by a gm-C filter as shown in Figure 1.2.30.

The core of the filter consists of two identical integrators (gm_2 with C_2 and gm_3 with C_1), which form a second-order resonator together with gm_1. The center frequency ω_0, the gain $Av[\omega_0]$, and the quality factor Q of the filter are defined by

$$\begin{aligned} \omega_0 &= \frac{gm}{C} = 5 \text{ kHz} \\ Q &= \frac{gm}{gm_1} = 5 \\ Av[\omega_0] &= Av[0]Q = \frac{gm_0}{gm_1} = 20 \\ gm_2 &= gm_3 = gm;\ C_1 = C_2 = C \end{aligned} \tag{1.2.34}$$

Fully differential design was used to achieve high common-mode and power-supply rejection ratios. A schematic diagram of the filter is shown in Figure 1.2.31.

The same transconductors that have already been used for the preamplifier are employed to achieve a wide linear range and the small transconductance needed for the low center frequency [43]. The most important parameters of the transconductors are summarized in Table 1.2.7.

Poly/poly capacitors were used to realize the integration capacitors. The capacitance per unit area of poly/poly capacitors is smaller than that of MOS capaci-

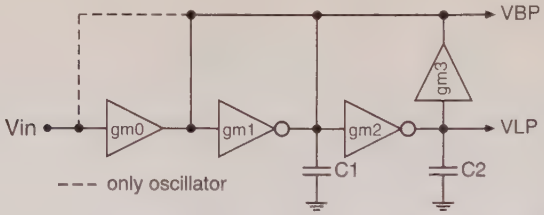


Figure 1.2.30. Block diagram of the gm-C bandpass filter and oscillator. The topologies of filter and oscillator are identical except for the feedback from the bandpass output to the input of *gm*0, which is only present in the oscillator (dotted line).

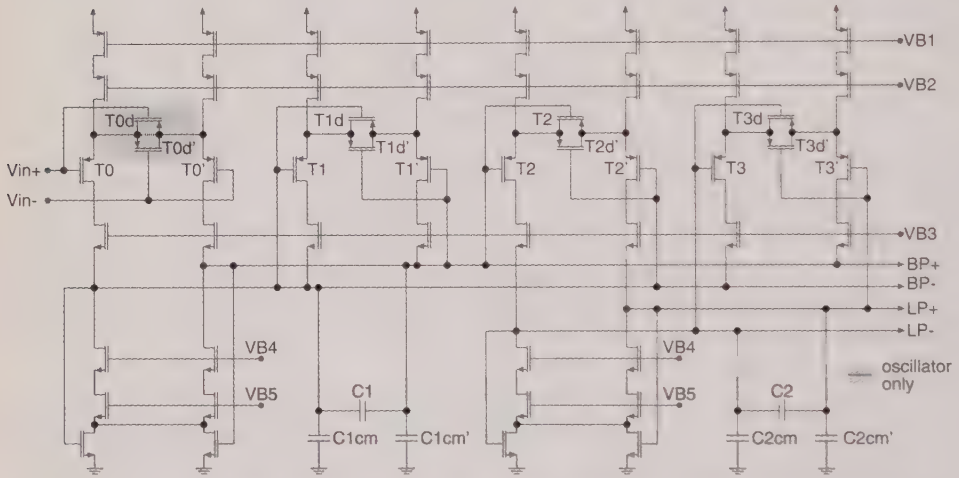


Figure 1.2.31. Schematic diagram of the fully differential gm-C filter used for the bandpass filter and the oscillator. For the oscillator, *gm*0 is formed by a simple differential pair instead of the degenerate transconductance. The biasing and CMFB are identical to the preamplifier.

Table 1.2.7. Summary of the design parameters of the transconductors used for the filter and the oscillator

Transistor	Bandpass filter			Oscillator		
	Size (μm/μm)	Current (μA)	*gm/a* (μA/V)	Size (μm/μm)	Current (μA)	*gm/a* (μA/V)
*T*0	96/20	10.0	16.0	24/20	0.125	1.6
*T*1	24/100	0.5	0.8	24/100	0.5	1.0
*T*2,3	24/20	2.5	4.0	24/20	2.5	4.0

tors but the use of floating instead of grounded capacitors reduces the size of the capacitor by a factor of four. Furthermore, poly/poly capacitors show better linearity. The floating capacitors also allow for independent design of the common-mode feedback. While the poles and zeros of the differential transfer function are set by the capacitors C_1 and C_2 (60 pF), the GBW of the CMFB is determined by C_{1CM} and C_{2CM}, respectively.

The circuit for the CMFB is needed twice: once to stabilize the lowpass output and once for the bandpass output. The same CMFB that was already used for the preamplifier in Section 1.2.3.5.4 is employed for the filter and oscillator. With a value of 2 pF for the capacitors C_{1CM} and C_{2CM}, the bandwidth of the two CMFBs is 200 kHz, which is a factor of 40 higher than the center frequency of the filter.

The noise performance is determined by the input transconductance gm_0, which is formed by the input transistors T_0 and T_{0d} and their biasing transistors. Together, they contribute two-thirds of the total noise. The corner frequency is 2 kHz, and the input-referred thermal noise floor is 90 nV/Hz$^{1/2}$.

The adjustable gain of 5/20 is realized by switching off 75% of the input transconductance in the case of the lower gain (not shown in Figure 1.2.31). The same technique was used for the preamplifier shown in Figure 1.2.26.

Chopping-clock Generator

The frequency of the oscillator used to generate the chopping clock must be accurately matched to the center frequency of the bandpass filter. In order to avoid clock feedthrough from the oscillator to the amplifier, the oscillator output runs at twice the chopping frequency and is subsequently divided by a flip-flop. The oscillator uses the same topology as the filter in order to achieve good matching (see Figure 1.2.30). The transconductances gm_1, gm_2, and gm_3 are exactly the same as those employed in the bandpass filter. The frequency is doubled by scaling down all the capacitors by a factor of two.

The oscillation amplitude is defined by the nonlinear transconductance formed by the parallel combination of gm_0 and gm_1 as shown in Figure 1.2.32. The di-

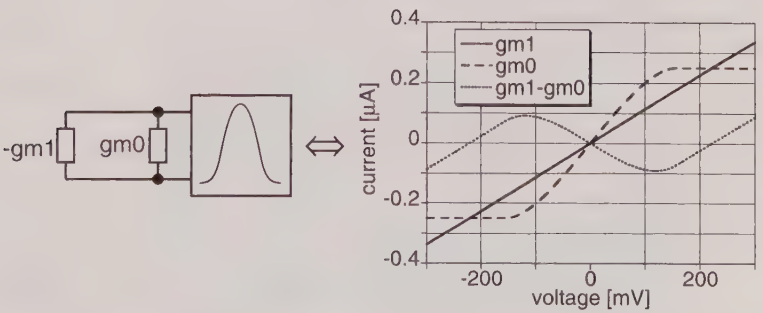


Figure 1.2.32. Schematic and simulated current–voltage characteristic of the parallel transconductance formed by *gm*0 and *gm*1.

mensions of transconductors gm_0 and gm_1 have been derived in [44]. For small input voltages, gm_0 dominates and leads to an increase in amplitude. For large amplitudes, gm_0 saturates and gm_1 reduces the amplitude towards a stable value.

The dimensions of the two transconductors are determined by the following considerations:

- The maximum input voltage for the simple differential pair formed by T_0 and T_0' is $2\times Vd_{sat}[T_0]$. A value of $\sim 5\times Vd_{sat}[T_1]$ is obtained for the degenerate transconductor gm_1 (see Figure 1.2.31). The oscillation amplitude must be chosen between these two voltages.
- A stable oscillation is obtained if the total energy dissipated in the transconductances during one clock cycle is zero:

$$E[V_{osc}] = \int_0^{T_0} V \cdot I \cdot dt = V_{osc}^2 \int_0^{T_0} \sin(2\pi f_0 t)^2 \cdot (gm_{0[t]} + gm_{1[t]}] \qquad (1.2.35)$$

- The slope of $E[V_{osc}]$ should have a maximum at $E[V_{osc}]=0$.

The oscillation amplitude was set to 540 mV_{pp}. Table 1.2.7 summarizes the parameters calculated for gm_0 and gm_1. The measured oscillation amplitude for 10 chips from different wafers was 556 mV with a standard deviation of 2.6 mV.

The topology of the chopping-clock generator is shown in Figure 1.2.25. In order to drive the chopper switches, the small sine-wave oscillation has to be converted into a square-wave signal. The conversion is done using a comparator with a small hysteresis of 40 mV as can be found in many textbooks (eg, [46]). A subsequent flip-flop serves as a frequency divider. In this way, a duty cycle of exactly 50% can be guaranteed. The comparator occupies an area of 0.1 mm^2 and consumes 25 µW from a 5 V supply.

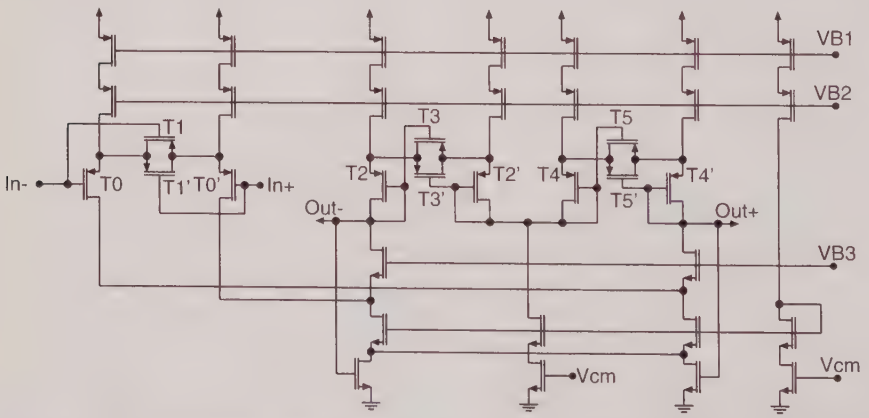


Figure 1.2.33. Schematic diagram of the third amplification stage of the chopper amplifier.

Table 1.2.8. Design parameters of the third amplifier

Gain	26 dB
Bandwidth	250 kHz
Power consumption	350 μW
Output swing	±3.2 V
Equivalent input noise	70 $nV/Hz^{1/2}$
Area	0.075 mm^2

1.2.3.5.6 Third Amplifier

The requirements of the third amplifier are low power, a gain of 20, an input swing of 100 mV, and an output swing of 2 V. Owing to the two preceding amplification stages, noise and offset performance of this amplifier are not critical. The design is shown in Figure 1.2.33.

The increased swing of the amplifier is realized by using a degenerate transconductance for the input stage and the splitting technique already employed for the weak inversion preamplifier. The most important parameters of the design are summarized in Table 1.2.8.

1.2.3.5.7 Measurement Results

Two test chips including only the chopper amplifier were designed to characterize the chopper amplifier; one uses the strong inversion preamplifier and the other the version operating in weak inversion. Figure 1.2.34 shows a micrograph of the chopper amplifier with strong inversion preamplifier. The size of the chip (excluding pads) is 1450×950 μm^2.

Table 1.2.9 summarizes the power consumption of the chopper amplifier. The measured parameters are summarized in Table 1.2.10. The gain and bandwidth of the three amplification stages are in good agreement with the designed values. The mismatch between the center frequency of the bandpass filter and the oscil-

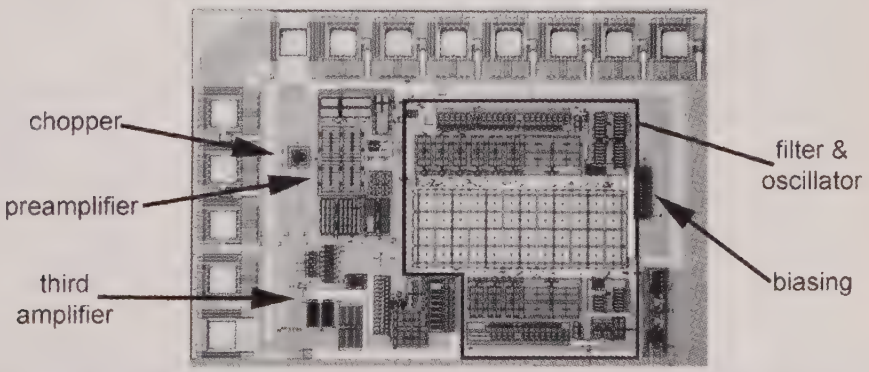


Figure 1.2.34. Micrograph of the test chip for the strong-inversion chopper amplifier.

Table 1.2.9. Power consumption of the chopper amplifier (mW)

Component	Weak inversion	Strong inversion
First amplifier	0.90	1.95
Bandpass filter	0.23	
Oscillator	0.15	
Third amplifier	0.35	
Total	1.63	2.58

Table 1.2.10. Summary of measured performance of the chopper amplifiers. Seven chips of each design were measured; for the weak inversion design they were taken from three different wafers

Parameter	Strong inversion		Weak inversion		Units
	Mean	S.D.	Mean	S.D.	
Gain	75.6	0.3	75.8	0.18	dB
Gain (tuned BP)	75.75	n.a.	75.9	n.a.	dB
Gain unchopped 5 kHz	77.6	0.06	77.8	0.06	dB
3 dB bandwidth	477	26	472	17.0	Hz
Equivalent input noise	7		8		$nV/Hz^{1/2}$
Input offset	0.24	0.6	0.1	0.86	μV
CMRR (0–500 Hz)	>140		>140		dB
Output range	±2.3	n.a.	±2.3	n.a.	V
Oscillator frequency	9812	121	9452	44.5	Hz
Bandpass center frequency	4948	88	4775	41.6	Hz
Matching	0.8	1.2	1.0	1.2	%

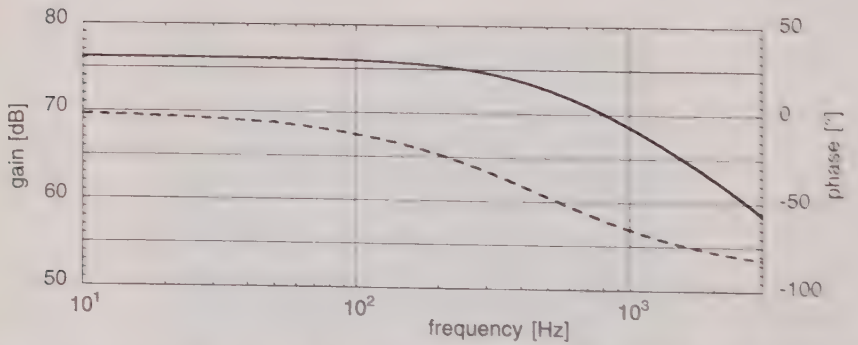


Figure 1.2.35. Gain-phase plot of the chopper amplifier after tuning of the bandpass center frequency.

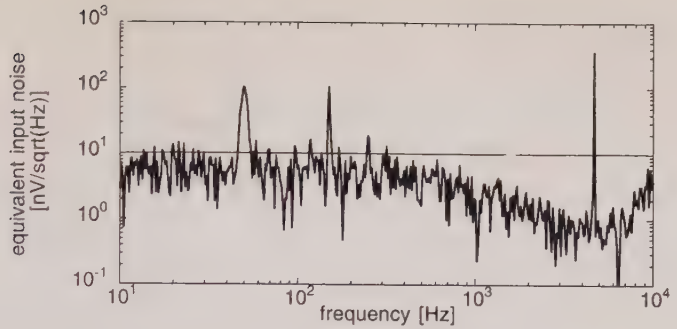


Figure 1.2.36. Equivalent input noise voltage of the chopper amplifier. The peaks at 50 and 150 Hz are due to line interference.

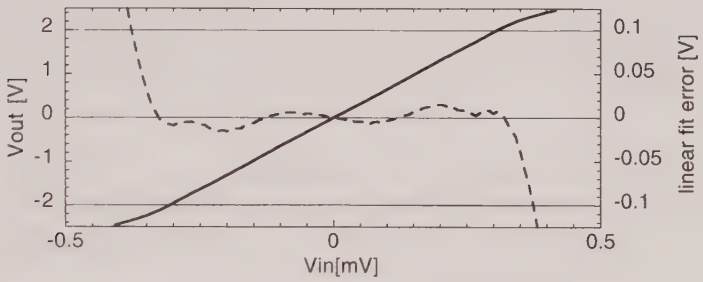


Figure 1.2.37. DC transfer characteristic and linearity error of the tuned chopper amplifier.

lation frequency is 1.2% for both designs, which leads to an average reduction of 0.1 dB in the DC gain of the chopper amplifier (see Equation (1.2.31)) and an increase in the residual offset (see Equation (1.2.33)). An external current to tune the center frequency of the bandpass filter can be applied for applications where this cannot be tolerated. In this way, the DC gain calculated in Equation (1.2.22) was exactly obtained. Figure 1.2.35 shows a Bode plot of the measured frequency response of the chopper amplifier. In Figure 1.2.36, the measured noise performance of the design is shown. The peaks at 50 and 150 Hz are due to line interference and were filtered for the calculation of the equivalent input noise. Figure 1.2.37 shows the DC transfer characteristic of the chopper amplifier when the maximum gain is selected. The nonlinearity is dominated by the degenerate transconductance of the third amplifier. This can be concluded from the fact that the amplitude and shape of the fit error are in good agreement with the measurements of the degenerate transconductance in [44].

The possibility of tuning the center frequency of the filter is not available for the complete multisensor chip, which includes all three sensors. The small variations in gain can be accepted because the sensitivity has to be calibrated in any case owing to the fabrication tolerances of the transducer. The offset is not criti-

cal at all, because the offset due to gas flow ($\sim 100\,\mu V$ at $200\,cm^3$) and other sources exceeds the offset of the amplifier by a factor of 100.

1.2.3.6 Antialiasing Filter and A/D Converter

The signal coming from the chopper amplifier needs to be converted to a digital signal with a resolution of 13 bits at a bandwidth of 400 Hz. Oversampled one-bit $\Sigma\Delta$ modulators offer two major advantages over other solutions:

- Elements that need a matching performance equal to the resolution of the A/D converter are not needed in a one-bit $\Sigma\Delta$ A/D converter.
- Nyquist rate A/D converters for low-frequency applications need complex high-order analog filters. These filters consume a large area owing to the large time constants. The oversampling relaxes the specifications of the antialiasing filter compared with Nyquist rate converters. The filtering of the out-of-band noise can be shifted to the digital domain, because the antialiasing filter is only needed to suppress noise at frequencies exceeding half the oversampling frequency.

The minimum oversampling ratio needed to achieve a resolution of 12 bits with a one-bit second order is 64 [47]. An oversampling ratio of 128 was chosen to relax the specifications of the decimation filter and the antialiasing filter. The result is a sampling frequency of 100 kHz for the $\Sigma\Delta$ modulator, which can be easily derived from the system clock of 800 kHz.

The antialiasing filter needs a total damping of 80 dB at the sampling frequency (minus the signal bandwidth, which can be neglected). In the case of a chopper amplifier with bandpass filtering, the noise spectrum at the output is not

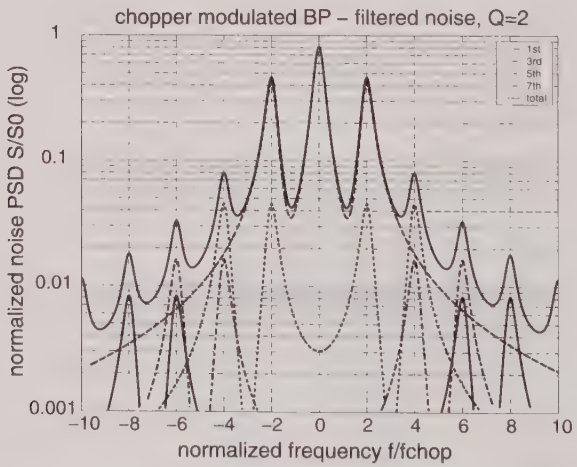


Figure 1.2.38. Noise spectrum at the output of the chopper amplifier. The noise is attenuated by 20 dB/decade due to the bandpass filter. Reprinted from [42].

white, but drops off with 20 dB/decade as can be seen in Figure 1.2.38 [42]. The attenuation of the antialiasing filter can be as low as 54 dB at 100 kHz. The same filter that has been used for the bandpass filter can also be used for the antialiasing filter, because it also offers a lowpass output (see Figure 1.2.30). Only two minor changes have to be implemented:

- gm_1 is made equal to $gm_{2,3}$ in order to reduce Q to one;
- gm_0 is also made equal to $gm_{2,3}$ in order to realize a DC gain of one;
- a small buffer (source follower) is needed at the output of the filter in order to drive the input sampling switches of the $\Sigma\Delta$ modulator.

With these changes, an attenuation of 53 dB is achieved at the sampling frequency.

The $\Sigma\Delta$ modulator is also re-used from another design: The interdigitated input and feedback capacitors of the $\Sigma\Delta$ modulator used for the capacitive sensor are simply replaced by standard poly/poly capacitors to transform it into a standard $\Sigma\Delta$ A/D converter. The reference voltage of 1.2 V is provided by a standard bandgap reference that is available from the library of the 0.8 µm process.

1.2.3.7 Decimation Filter

The digital filter has four main tasks:

- Filtering of the shaped high-frequency quantization noise of the $\Sigma\Delta$ converter.
- Limiting the bandwidth of the signal to 400 Hz. This removes the thermal noise floor at frequencies higher than the bandwidth, and filters the modulated $1/f$ noise from the third amplifier.
- Filtering the chopped DC offset of the bandpass-filter in the chopper amplifier.
- Downsampling of the bitstream coming from the converter by a factor of 128 and delivery of a 12-bit word at the output.

The difficult task in designing hardwired decimation filters in CMOS processes is finding a topology that can achieve the desired performance for minimum power consumption and area. Many architectures (eg, [48]) that have been proposed for digital signal processors (DSPs) in the literature are not well suited for hardwired filters, because they use multipliers. Multipliers consume a lot of area, if accuracies of 12 bits and better are required. Filters based on multiplier-free topologies [49, 50] and comb filters are preferred in order to achieve an area-efficient solution. Candy [51] showed that a simple sinc^l-filter can be used for decimation down to four times the output-word rate f_0 without degrading the SNR. The order l of the filter must equal the order of the $\Sigma\Delta$ modulator plus one. For the subsequent decimation to f_0, a finite-impulse response (FIR) filter with a

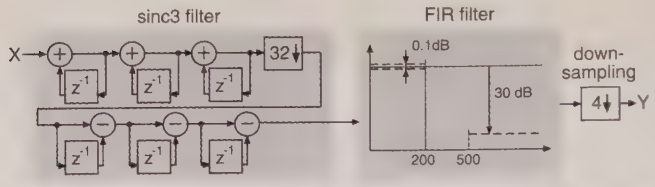


Figure 1.2.39. Decimation filter with downsampling factor of 128 and an output word length of 13 bits. The SNR after the filter is 73.1 dB, taking into account quantization noise, thermal noise of the $\Sigma\Delta$ modulator, and the effects of truncation through the digital filter.

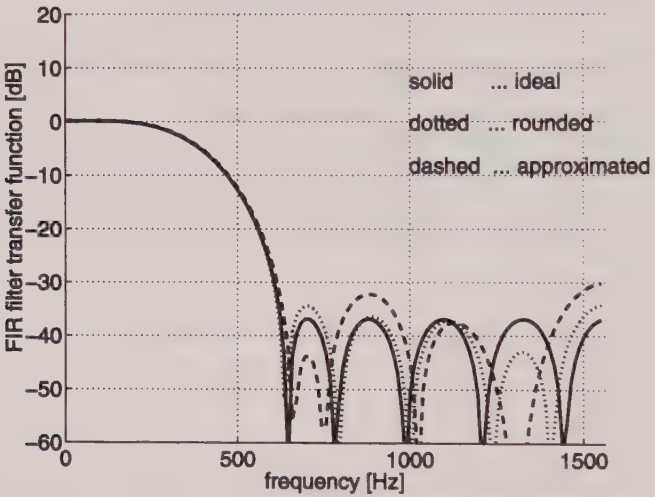


Figure 1.2.40. Transfer function of the 13-tap FIR filter. The solid line represents the ideal transfer function with the coefficients calculated by the Remez exchange algorithm. The dotted and dashed lines show the effects of rounding the coefficients to 8 bits and approximating them with the sum/difference of two power-of-two numbers.

higher order cut-off is required. Figure 1.2.39 shows the architecture of the combined decimation filter and the frequency specifications of the FIR filter. A maximum passband ripple of 0.1 dB (passband corner at 200 Hz) and a minimum stopband attenuation of 30 dB were specified.

Starting from the loss characteristics shown in Figure 1.2.39, the coefficients of the FIR filter were calculated using the Remez exchange algorithm [52]. The coefficients were then approximated by the sum of two power-of-two coefficients. This allows for replacement of all multipliers in the system by two shifters and an adder:

$$k[i] \approx \pm 2^n \pm 2^m \tag{1.2.36}$$

The result was a 13-tap filter with a word length of 8 bits for each coefficient. Figure 1.2.40 shows the ideal transfer function and the effects of rounding the coefficients to 8 bits and approximating them by power-of-two coefficients.

Table 1.2.11. Design parameters of the $sinc^3$ filter

Word length of integrators			Word length of differentiators		
No. I	No. II	No. III	No. I	No. II	No. III
16	15	14	12	12	12

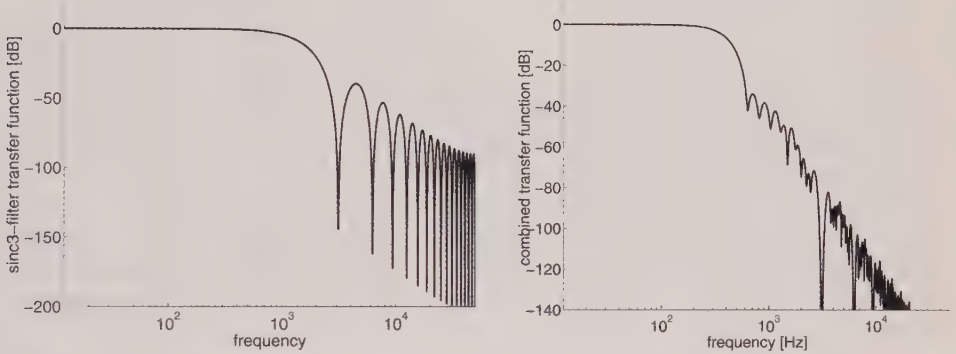


Figure 1.2.41. Transfer function of the $sinc^3$ filter and combined transfer function of the two digital filters.

The word lengths of all the registers and the effects due to finite word length in the filter were then minimized by calculating the SNR for different configurations. The implementation with the smallest overall area, which still guarantees a resolution of 12 bits (SNR = 74 dB), had to be found. The optimization was done using the signal processing and discrete time toolboxes of Matlab™ [53]. The results for the $sinc^3$ filter are summarized in Table 1.2.11. The SNR at the output of the filter was simulated to be 74.3 dB.

The filter was then described in VHDL and automatically synthesized using Synopsys™. The filter occupies a chip-area of 4 mm^2. Further reduction of the filter area is possible if bit-serial architectures or topologies with shared adders are used. This is feasible since, eg, the FIR filter operates at a clock-frequency that is 256 times lower than the system clock.

1.2.3.8 On-chip Calibration

The physical sensitivity of the calorimetric sensor shows large variations due to processing tolerances of the sensitive layer, the membrane, the thermocouples, and the read-out circuitry. Calibration is necessary but no accurate reference is available. On-chip calibration can be performed by placing a heating resistor in the center of the membrane and applying a defined heating power to this resistor.

The heating current is a multiple of the global bias current, which is derived from a bandgap voltage and hence well defined.

1.2.4 Resonant-beam Gas Sensor

1.2.4.1 Sensing Principle

The analyte-induced mass uptake of a sensitive polymer can be used to classify different gases or to quantify a known analyte. The different miniaturized devices to assess the mass change of a sensitive layer (mostly polymeric) upon analyte absorption from the gas-phase have been discussed [1]. The operation principle of the resonant cantilever which is used for the multisensor chip is shown in Figure 1.2.42. Upon absorption of an analyte into the polymer on the silicon cantilever, the mass load is increased and, therefore, the resonance frequency of the system decreases. The sensitivity S of a silicon cantilever coated with a thin polymer layer has been derived [54]:

$$\begin{aligned} S &= \frac{\partial f_0}{\partial \rho_L} \cdot \frac{\partial \rho_L}{\partial c_A} \\ &= G \cdot S_A \end{aligned} \tag{1.2.37}$$

where f_0 denotes the resonance frequency of the cantilever and c_A is the analyte concentration in the gas phase. Equation (1.2.37) shows that the resonance frequency change is due to a polymer density change $\delta\rho_L$ upon analyte absorption. S_A relates to the polymer–analyte interactions and equals the partition coefficient K_c (mass/volume) when mass concentrations are used. G represents the mechanical properties of the cantilever device and is given by

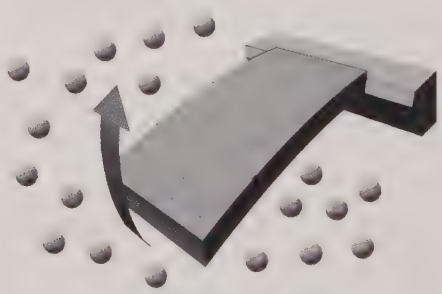


Figure 1.2.42. Polymer-coated resonant silicon cantilever.

$$G = \frac{\partial f_0}{\partial \rho_L} = \frac{\partial f_0}{\partial \rho_{mean}} \cdot \frac{\partial \rho_{mean}}{\partial \rho_L} = \frac{1}{2} \frac{f_0}{\rho_{mean}} \frac{t_L}{h} \tag{1.2.38}$$

A detailed explanation of G can be found in [54]. ρ_{mean} denotes the average specific mass density of the cantilever layer stack (silicon, oxide, nitride, polymer), t_L is the thickness of the polymer and h is the total thickness of the cantilever.

Finally, the sensitivity S [Hz/(µg/l)] of the cantilever gas sensor can be rewritten as

$$S = -\frac{1}{2} \frac{f_0}{\rho_{mean}} \frac{t_L}{h} \cdot K_c \tag{1.2.39}$$

1.2.4.2 Thermomechanical Actuation and Piezoresistive Detection

1.2.4.2.1 Feedback Operation

There are basically three ways to determine the resonance frequency of a cantilever:

1. Noise spectrum: The frequency spectrum of the thermal noise is measured and analyzed. The peak frequency is equivalent to the resonance frequency. This method does not require any excitation of the cantilever, but a low-noise spectrum analyzer is needed for the measurement.

2. Frequency sweep of excitation signal: The cantilever is excited at different frequencies and the response is evaluated to determine the resonance frequency. A gain-phase meter is needed for the measurement.

3. Oscillator: The cantilever is used as the loop-filter in an oscillator. A fast feedback is needed in this case but a simple counter can be used to determine the resonance frequency.

The third option leads to the most accurate results because the Q-factor of the cantilever is enhanced by the feedback. An actuation mechanism and a deflection sensor are needed to realize a monolithically integrated design of an oscillator. Furthermore, only the third option can be realized in CMOS technology with reasonable effort, because integrated spectrum analyzers and gain-phase meters are difficult to design.

1.2.4.2.2 Cantilever Design

The design considerations for cantilever beam resonators have been described in detail [55]. Figure 1.2.43 shows a micrograph of the cantilever that has been optimized for gas-sensing applications. The heating resistors for thermomechanical actuation (Section 1.2.4.2.4) and the piezoresistors, which detect the deflection of the beam, are highlighted in black. The most important features of the cantilever are summarized in Table 1.2.12.

1.2.4.2.3 Detection of the Cantilever Deflection

In desktop systems, a high-precision optical read-out scheme using a laser beam is often used to measure the deflection of the cantilever. In the integrated approach, the deflection of the cantilever is frequently assessed by measuring the stress induced at the clamped edge of the cantilever. This stress is measured using the piezoresistive effect of a material that is ready available in the CMOS process (eg, monocrystalline silicon) [56].

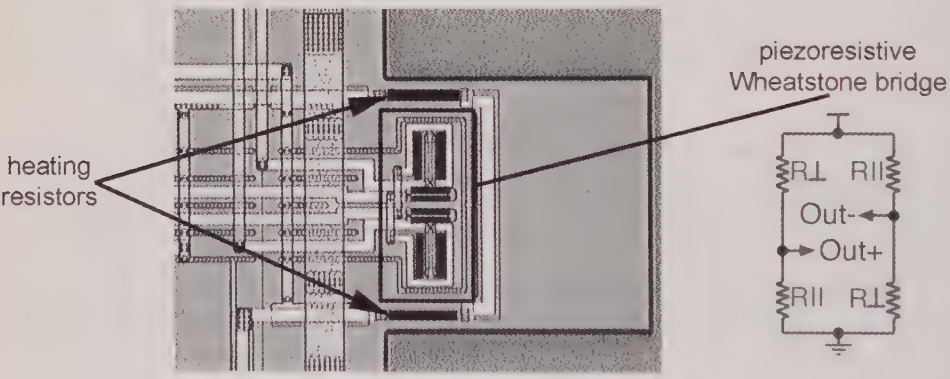


Figure 1.2.43. Cantilever with diffused heating resistors and piezoresistive Wheatstone bridge. The resonance frequency is 380 kHz ± 5%.

Table 1.2.12. Summary of important features of the resonant cantilever

Parameter	Value
Beam size	$150\times140\times5\ \mu m^3$
Actuator	Diffused heating resistor (750 Ω)
Deflection sensor	Piezoresistive Wheatstone bridge (850 Ω)
Q-factor (in air)	950
Resonance frequency	380±20 kHz
Sensitivity	0.11 V/W at $V_{Bias}=5$ V, $f=f_{res}$

Piezoresistive Detection

The piezoresistive effect on a low p-doped silicon resistor is described by the relative change of resistance $\Delta R/R$:

$$\frac{\Delta R}{R} = \pi_L \sigma_L + \pi_T \sigma_T = 71.8 \cdot 10^{-11}\ \text{Pa}^{-1} \cdot \sigma_L - 66.3 \cdot 10^{-11}\ \text{Pa}^{-1} \cdot \sigma_T \quad (1.2.40)$$

where $\pi_{L,T}$ are the longitudinal and transversal piezoresistive coefficients and $\sigma_{L,T}$ denote the respective stress components. According to Equation (1.2.40), the piezoresistive coefficients parallel and perpendicular with respect to the cantilever axis have opposite signs. Therefore, a differential signal is obtained by arranging two resistors parallel and two resistors perpendicular to the cantilever axis in a Wheatstone bridge configuration (see Figure 1.2.43). The common-mode voltage varies by only a few percent because the absolute values of the longitudinal and transversal piezoresistive coefficients are almost identical.

The resistors in the Wheatstone bridge have a value of 850 Ω. This leads to a power-consumption of 29 mW. As minimum width resistors had to be chosen for the Wheatstone bridge owing to area restrictions, a good matching cannot be expected. The matching of the resistors is further deteriorated owing to their perpendicular orientation. This leads to a large offset voltage of up to 20 mV.

Diffused resistors were used in designing the single-chip gas sensor system. The following paragraphs discuss the advantages and disadvantages of two alternative approaches for integrated deflection measurement.

Stress-sensitive MOS Transistor

Mechanical stress changes the carrier mobility in the channel of MOS transistors [55]. Four diode-connected MOS transistors in a bridge configuration are used to generate an output voltage proportional to the stress at the clamped edge of the cantilever [60] (like in a full Wheatstone bridge, opposite signs of the stress-sensitivities are achieved by orienting the channels of the two pairs of transistors perpendicular to each other).

The diode-connected MOS transistors show a smaller small-signal transconductance than diffused resistors, which reduces the power and area consumption. The disadvantages are increased noise and a reduction in the sensitivity by approximately 50%.

The area of the bridge is only critical for designs with resonance frequencies in excess of 1 MHz. The single-chip gas sensor system uses a cantilever with a resonance frequency of 380 kHz and the power consumption was not the main optimization criterion. In this case, the piezoresistive Wheatstone bridge is the preferred solution for stress detection.

Piezoelectric Detection

The advantage of piezoelectric detection is a self-generated signal, which eliminates the need for biasing and reduces the offset. CMOS processes do not in-

clude a layer of piezoelectric material. The CMOS-compatible deposition of piezoelectric material is difficult [57], adds extra post-processing steps, and is hard to combine with the deposition of the sensitive layer.

1.2.4.2.4 Thermomechanical Actuation

Different methods have been developed to excite a cantilever. Capacitive actuation is frequently used, eg, for acceleration sensors [79]. It is not suitable for gas-sensing applications, because it requires a counter electrode close to the cantilever. The fabrication of the counter electrode is only possible with additional post-processing steps. Furthermore, the counter electrode renders the coating of the cantilever difficult and causes problems when the device is exposed to a gas stream (dust, sticking). The disadvantages of piezoelectric actuators have already been described. In addition to the deposition problems, piezoelectric actuation is inefficient and causes considerable heating of the cantilever. This reduces the sensitivity of the gas sensor owing to a reduction in the partition coefficient (see Section 1.2.2.3). Both techniques make it difficult to deposit a sensitive layer. In some cases, electromagnetic actuation is a promising approach [55].

Thermomechanical actuation using the bimorph effect between the silicon n-well and the dielectric layers was found to be the best solution for the multisensor chip. This actuation mechanism has already been used to excite silicon membranes for ultrasound generation [58] or silicon beams in acceleration sensors [59]. Compared with other oscillators (eg, quartz crystal, LC tank) a few difficulties have to be taken into account when a thermomechanically actuated cantilever is used as the frequency-determining element.

Efficiency of the Thermomechanical Actuation

The transfer function of the cantilever (without thermal or capacitive crosstalk) shown in Figure 1.2.44 has been calculated [55]. Owing to the thermomechanical

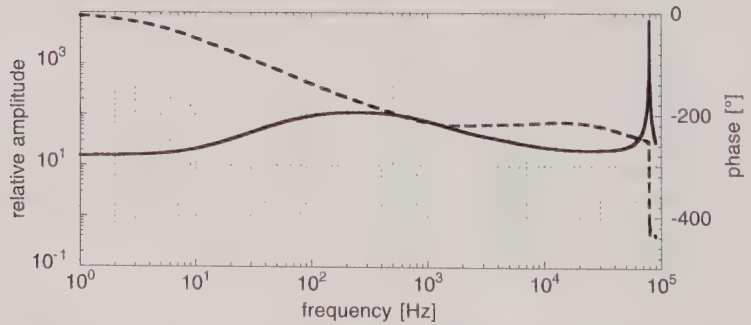


Figure 1.2.44. Calculated amplitude (solid line) and phase lag (dashed line) of a thermomechanically excited cantilever beam as a function of frequency.

excitation it looks completely different to an LC tank or quartz oscillator. With increasing frequency, the thermal wave is more and more confined around the heating resistor. The efficiency of the thermomechanical excitation decreases. This reduces the amplitude of the oscillation and leads to a phase lag that increases with frequency.

Influence of Polymer Coating

Figure 1.2.45 shows a Bode plot for a cantilever before and after coating with PEUT. After coating, the resonance frequency and the amplitude at resonance are decreased. A larger loop gain is needed to compensate for the decrease in amplitude. For a general harmonic oscillator, the phase shift at resonance between excitation and response is independent of the resonance frequency. Owing to the thermomechanical actuation, the resonant cantilever shows an additional phase lag, which increases with frequency as shown in Figure 1.2.44. Therefore, the phase shift at resonance is decreased after the coating procedure and depends on the actual thickness of the polymer. Owing to this decrease and the fabrication tolerance-induced fluctuation of the initial resonance frequency (<5%), the phase at resonance cannot be predicted accurately. The feedback circuitry must therefore include an adjustable phase shifter in order to guarantee a positive feedback.

The Q-factor of the cantilever and accordingly the stability also decrease with increasing polymer thickness. While the sensitivity increases with polymer thickness, the minimum detectable frequency shift is defined by the stability of the oscillator. This leads to a trade-off for the polymer thickness. For PEUT, an optimum thickness of $\sim 6\,\mu m$ was found [55]. The resulting shift in resonance frequency is 50 kHz and the Q-factor is decreased to from 950 to 600.

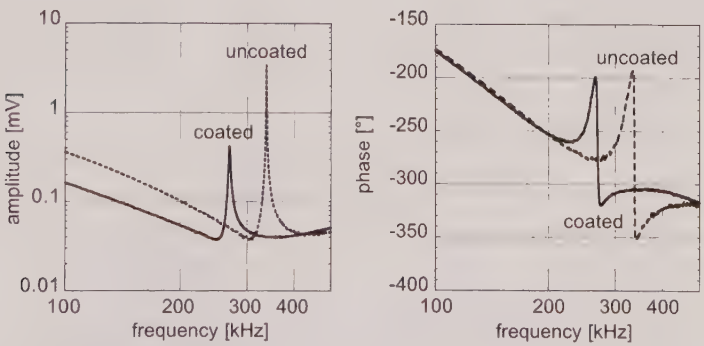


Figure 1.2.45. Measured amplitude and phase of the same cantilever (after amplification by the DDA, 30 dB) before and after coating with PEUT. The amplitude and phase are dominated by the mechanical signal only in the vicinity of the resonance frequency. At lower and higher frequencies, thermal and capacitive crosstalk is dominant (see Section 1.2.4.2.5).

Deflection Proportional to Heating Power

The deflection of the cantilever is proportional to the applied heating power. Therefore, a sine-wave excitation voltage cannot be used because all the heating power is then located at DC and twice the resonance frequency. A DC offset has to be added in order to excite the fundamental resonance of the cantilever:

$$\begin{aligned} P_{\text{heat}} &= \frac{[U_{\text{heat}} \sin(\omega t) + U_{\text{DC}}]^2}{R_{\text{heat}}} \\ &= \frac{\frac{1}{2} U_{\text{heat}}^2 + U_{\text{DC}}^2}{R_{\text{heat}}} + \frac{2 U_{\text{heat}} U_{\text{DC}} \sin(\omega t)}{R_{\text{heat}}} - \frac{U_{\text{heat}}^2 \cos(2\omega t)}{2 R_{\text{heat}}} \end{aligned} \tag{1.2.41}$$

The frequency component at the excitation frequency is proportional to the DC voltage and hence equal to zero, if no offset voltage is applied. An adjustable phase shifter, which corrects the frequency dependence of the phase (see above), is difficult to realize. If the signal is transformed into a square wave by a comparator, then the phase shift can be realized by a tunable delay line that can easily be implemented. The square-wave signal also solves the problem of the DC offset and no gain control is needed to achieve constant oscillation amplitude.

Small Signal Amplitudes

The signal from the Wheatstone bridge of an uncoated cantilever is only 1.8 mV_{pp} for a square-wave excitation signal of 3.5 V. After coating, the amplitude is reduced by 40% to 1.08 mV_{pp}. Therefore, a minimum loop gain of 1650 at the resonance frequency is needed to achieve stable oscillation.

1.2.4.2.5 Thermal and Capacitive Crosstalk

In Figure 1.2.46, the spectral response of the Wheatstone bridge output of a single cantilever is compared with optical measurements. It is found that the signal

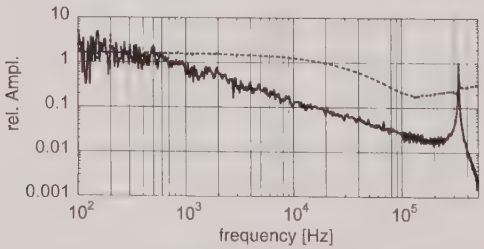


Figure 1.2.46. Spectral response of thermomechanically excited cantilever measured optically (full line) and at the output of the piezoresistive Wheatstone bridge (broken line). At low frequencies, the signal from the Wheatstone bridge is dominated by thermal crosstalk. Capacitive crosstalk is observed for frequencies above 200 kHz.

from the mechanical vibration dominates only in the vicinity of the resonance frequency. At low frequencies, the thermomechanical actuation heats the whole cantilever and creates a temperature gradient in the Wheatstone bridge. Owing to the large temperature coefficient of diffused resistors, the resulting signal is larger than the mechanical response for frequencies up to 200 kHz (for an uncoated cantilever). Besides the offset of the Wheatstone bridge, this is a second reason to eliminate the low-frequency signals in the feedback circuit, since the amplitude of the thermal crosstalk almost exceeds the amplitude at resonance. For frequencies higher than 200 kHz, capacitive crosstalk is observed. Owing to the small distance between the diffused heating resistors and the diffused resistors of the Wheatstone bridge, the parasitic capacitances through the n-well cannot be neglected.

For the design of the multisensor chip, the layout of the cantilever was optimized using grounded diffusion stripes in order to minimize the parasitic capacitances. With the improved design, the amplitude of the capacitive crosstalk does not exceed the amplitude of the mechanical response for frequencies up to 1 MHz (see Section 1.2.4.3.2).

1.2.4.3 Resonant-beam Oscillator

1.2.4.3.1 Feedback Topology

From the above considerations, the circuitry for the oscillator as shown in Figure 1.2.47 can be derived. The output signal of the Wheatstone bridge is first amplified by a low-noise differential difference amplifier (DDA). The amplification factor can have a maximum value of 35 in order to avoid saturation of the amplifier by the DC offset of the Wheatstone bridge. The signal is then high-pass filtered to remove the offset voltages of the bridge and the first amplifier. The high-pass filter also prevents upconversion of the amplifiers $1/f$ noise. Finally, the filter eliminates the low-frequency thermal crosstalk and therefore avoids resonances due to this mechanism. AC coupling at the input of the first amplifier would allow for a higher gain in the first amplification stage. The disadvantages are added noise and an additional path for switching interference and substrate noise to the most critical point in the circuit.

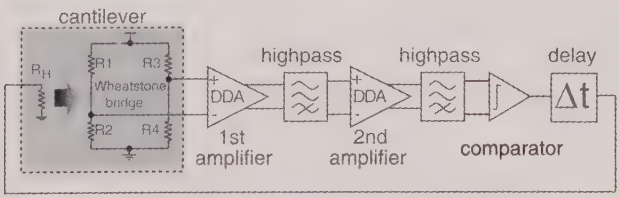


Figure 1.2.47. Oscillator with thermomechanically actuated cantilever as frequency-determining element.

The signal is then amplified and high-pass filtered a second time before it is converted into a square-wave signal by the comparator. The second amplification stage is needed to achieve a sufficiently large amplitude at the input of the comparator. The minimum amplitude at the input of the comparator is defined by

- The input offset of the comparator: an amplitude at least 10 times larger than the offset-voltage ($\sim$1 mV) is needed for the desired duty cycle of 45–55%.
- The noise and crosstalk at the input of the comparator: switching of the comparator due to noise coupling from the digital circuitry and the large switching transistor prevents stable oscillation. A hysteresis at the input of the comparator cannot be used because this would prevent the onset of the oscillation due to thermal noise in the Wheatstone bridge and mechanical noise from the cantilever.

After the comparator, a digital delay line is used to adjust the phase in order to achieve positive feedback at the resonance frequency. A source follower at the output of the delay line is used to drive the small heating resistor. In the following section, the building blocks of the feedback are described in more detail.

1.2.4.3.2 Differential Difference Amplifier (DDA)

The most important requirements for the first amplifier are low noise, a gain of 30, a bandwidth of 500 kHz, and a common mode rejection ratio (CMRR) of at least 80 dB at the resonance frequency. The DDA for instrumentation applications was proposed in the 1970s [60], and a version with differential output has been proposed [61]. In the configuration shown in Figure 1.2.48, the DDA is used as a fully differential instrumentation amplifier. Compared with the classical instrumentation amplifier configuration based on two operational amplifiers, the DDA has important advantages:

- The area and power consumption are lower because only one extra input stage (and a common-mode feedback) is added to a conventional operational amplifier.

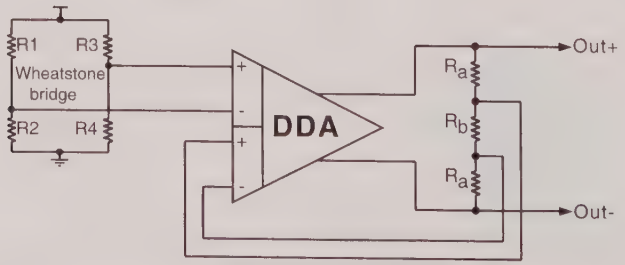


Figure 1.2.48. Differential difference amplifier used as an instrumentation amplifier.

- The common mode feedback (CMFB) of the DDA suppresses common-mode variations at its input whereas the classical configuration is transparent for common-mode signals. This relaxes the CMRR specifications of the subsequent stages.
- The common-mode voltage at the output of the DDA can be chosen regardless of the input common-mode voltage.

The general disadvantages of the DDA include its limited linearity and input range. They are not important in this application because of the small input signal from the Wheatstone bridge and the limiter after the second amplification stage.

The gain A_V of the first amplification stage is defined by the feedback resistors R_a and R_b:

$$A_V = 1 + \frac{R_b}{2R_a} \tag{1.2.42}$$

Figure 1.2.49 shows a schematic diagram of the DDA. A folded cascode configuration was chosen for the first amplification stage in order to achieve a maximum input and output range. The second amplification stage is needed to drive the feedback resistors. The resistors are made from the low-resistive gate-poly, because no high-resistive poly is available in the CMOS process used for this design. A Miller capacitor with a series resistor was used for pole splitting and compensation of the positive zero. For the CMFB, a simple differential stage with a diode-connected load transistor was used to bias transistors $T2$ and $T2'$. The common-mode voltage at the output is derived by splitting resistor R_a in Figure 1.2.48 into two parts. The output stage (transistor $T7$) is biased separately to enable measurements under different bias currents without changing the output characteristics of the amplifier.

The sizes and bias currents of the most important transistors are summarized in Table 1.2.13. The four input transistors have a common centroid layout to im-

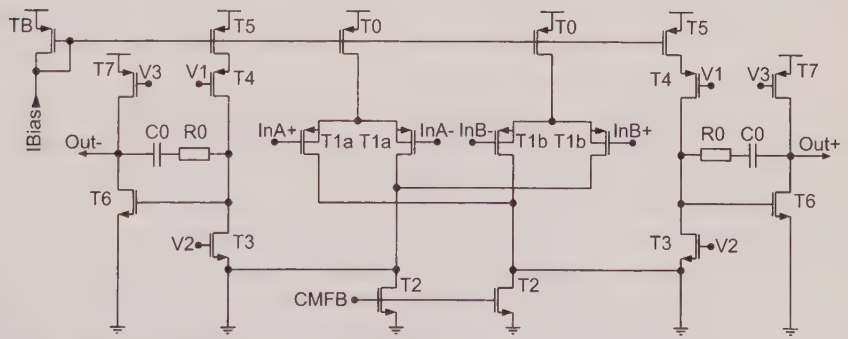


Figure 1.2.49. Schematic diagram of the differential difference amplifier.

Table 1.2.13. Transistor sizes and bias currents used for the DDA

Transistor	W (μm)/L (μm)	Current source	Current (μA)
$T1$	600/2	$T0$	2×100
$T2$	40/3	$T5$	35
$T3$	200/2	$T7$	250
$T4$	15/2		
$T5$	25/5		
$T6$	70/2		
$T7$	160/1		

Table 1.2.14. Design parameters of the DDA

Parameter	Simulation	Measurement
GBW (MHz)	18.5	21.9±0.75
Phase margin (°)	89	
GBW_{CMFB} (MHz)	19.5	
Phase margin CMFB (°)	52	
Open loop gain (dB)	118	
Closed loop gain (dB)	30.3	31±0.268
Power consumption (mW)	6	6
CMRR at 350 kHz (dB)		80
Input range (mV)	±50	±50
Common-mode input range (V)	2.5±1.6	2.5±1.6
Equivalent input offset (mV)		–0.563±1.1

prove matching of the two input stages, and they consume 35% of the total area of 0.175 mm^2. Feedback resistors with a total resistance of 15 kΩ are needed to achieve acceptable power consumption. These resistors account for another 12% of the total area.

Simulation and Measurement Results

Table 1.2.14 lists the design parameters of the DDA and compares the values obtained from simulation with the measured results. The measurements were done with 10 chips taken from different wafers of the same fabrication lot. The total harmonic distortion for realistic input signals <1 mV is 0.6%.

Figure 1.2.50 shows a measured Bode plot of the DDA transfer function. The bandwidth is 500 kHz as designed and the phase shift caused by the DDA at the cantilever resonance (380 kHz) is 35°. The measured thermal noise floor of 10 $nV/Hz^{1/2}$ is in good agreement with the designed value of 14 $nV/Hz^{1/2}$ and the $1/f$ corner frequency is 100 kHz. The noise performance is determined only

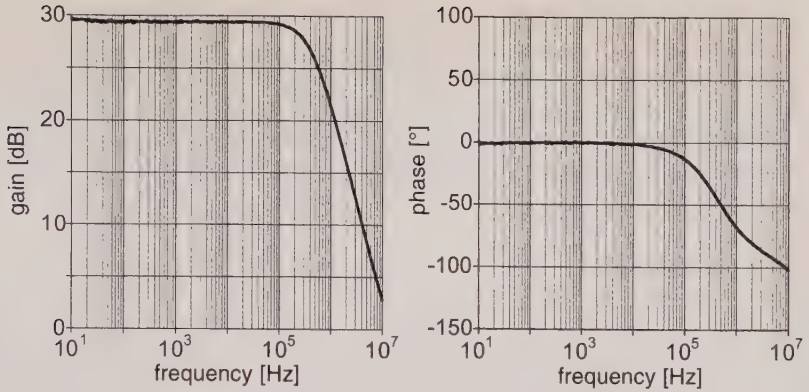


Figure 1.2.50. Bode plot of the transfer function of the DDA.

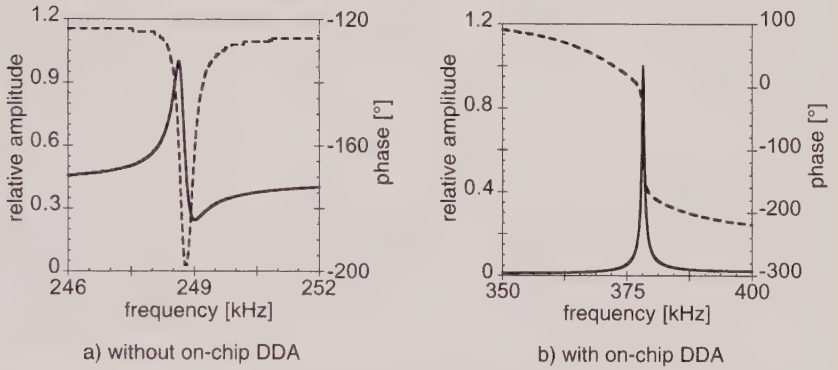


Figure 1.2.51. Gain-phase plot of the transfer function of the cantilever with and without on-chip DDA. The DDA reduces the amplitude of the capacitive crosstalk (solid line) by two orders of magnitude. The phase (dashed line) shows a clear phase-drop of 180° with the DDA.

by the input transistors $T1$ (46%) and the NMOS biasing transistors $T2$ (47%). All other transistors account for only 7% of the noise power.

Performance Improvement Through On-chip Circuitry

In Figure 1.2.51 the performance of a cantilever with on-chip DDA is compared with the performance of a chip without on-chip circuitry [62]. Both designs were packaged in standard DIL-28 packages and evaluated with a spectrum analyzer. Without an on-chip amplifier, high crosstalk due to the parasitic capacitances of the bond pads and instrumentation as well as the inductance of the bond wires is observed. The amplitude at resonance shown in Figure 1.2.51a (solid line) exceeds the crosstalk by a small factor of two. The phase (dotted line) shows a

drop of only 80° around the resonance frequency and is mostly determined by the crosstalk signal. This makes it very difficult to incorporate this cantilever into an oscillator. The measurement in Figure 1.2.51 b was recorded with on-chip DDA. The ratio between signal and crosstalk is now 100 and a clear phase drop of 180° around the resonance is observed. The additional phase lag is caused by the thermomechanical actuation (see Section 1.2.4.2.4).

1.2.4.3.3 Delay Line

The delay line is needed to compensate the phase shifts caused by the thermal excitation and the amplifiers in order to achieve positive feedback. As no external adjustment is available on the multisensor chip, only a digitally programmable delay using discrete delay increments can be implemented. From measurements of the frequency stability versus the phase shift it was found that a very coarse adjustment with a step size of 40° is sufficient to achieve the maximum stability [55]. A phase shift of 180° corresponds to a delay of 1.25 ms assuming a cantilever resonance frequency of 400 kHz. This delay is too large to be realized using digital elements such as buffers or inverters. Therefore, a delay line consisting of four unit-delay elements as is shown in Figure 1.2.52 was chosen. A 180° phase shift is implemented by simply using an inverter that can be switched into or out of the signal path. A schematic diagram of the unit-delay circuit is shown in the lower part of Figure 1.2.52. After a buffer, a constant bias current is used to charge/discharge a capacitor to the switching threshold of a subsequent Schmitt trigger. The hysteresis of the Schmitt trigger is needed to avoid bouncing of the signal due to substrate noise or other interferences as long as the voltage on the capacitor is close to the threshold voltage. A demultiplexer is used to select the right delay time. The delay is given by

$$t_{\text{delay}} = n \cdot \Delta t = n \cdot \frac{V_{\text{threshold}}}{I_{\text{Bias}} C} \tag{1.2.43}$$

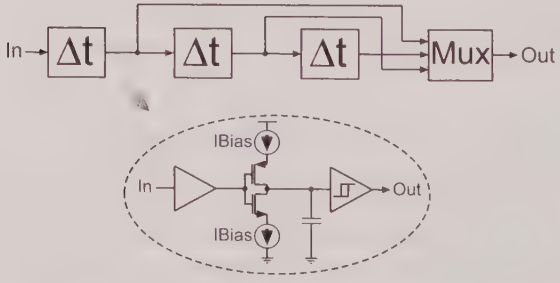


Figure 1.2.52. Block diagram of the delay line used to adjust the phase of the heating current in order to achieve positive feedback. This is done by selecting multiples of the unity delay Δt.

where n is the selected input of the multiplexer. The process tolerances of the cantilevers, the changes of the resonance frequency due to the application of the sensitive layer, and the tolerances of the resistors and capacitors make it necessary to implement a coarse tuning of the unit-delay by adjusting the current I_{Bias}. This is done using a programmable current mirror. The nominal value of 25 μA can be divided by a factor of 2, 3, or 4 to account for tolerances.

1.2.4.3.4 High-pass Filter

The only requirement for the high-pass filter is a corner frequency that is at least three times smaller than the resonance frequency of the cantilever. Therefore, a simple first-order RC filter can be used (see Figure 1.2.53). A value of 15 pF was chosen for the poly/poly capacitor C_0. This leads to a minimum resistor value of 3 MΩ taking into account the process spread. A transistor biased in the linear region was employed, because this large value cannot be realized with poly resistors. Corner simulations were performed to find the appropriate W/L ratio of 6 μm/150 μm at a bias voltage of V_{GS}=–2.5 V.

1.2.4.3.5 Measurement Results

The multisensor chip does not provide access to the analog signals of the resonant-beam oscillator. Therefore, a test chip containing only the resonant cantilever and its feedback circuitry was fabricated in order to characterize the performance of the sensor (see Figure 1.2.54). Figure 1.2.55 shows the measured

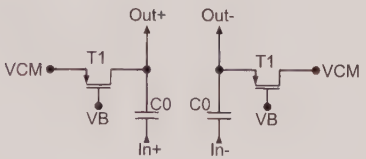


Figure 1.2.53. First-order high-pass filter needed for DC rejection and filtering of the 1/*f* noise from the DDA.

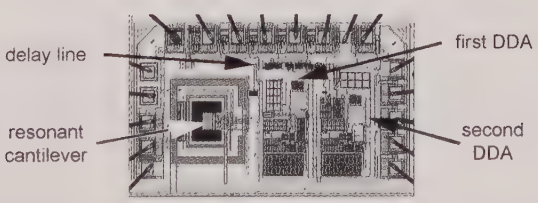


Figure 1.2.54. Micrograph of the test chip for characterization of the resonant-beam oscillator.

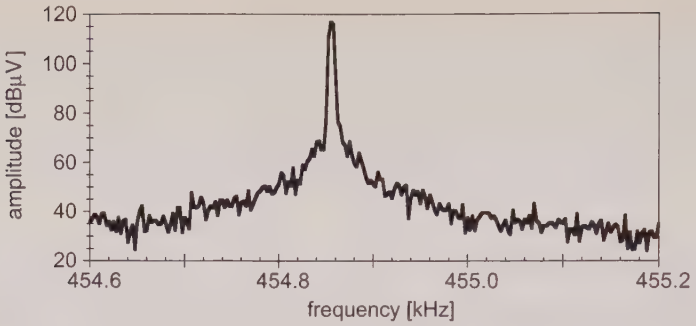


Figure 1.2.55. Measured spectrum of the resonant beam oscillator. The resonance frequency of the cantilever is at 455 kHz.

spectrum of the oscillator. The thermal noise floor is 80 dB below the oscillation amplitude and is a result of the thermal noise of the first amplifier and the Wheatstone bridge. At offset frequencies below 100 Hz from the resonance frequency, the upconverted thermal noise with its $1/f^2$ frequency behavior can be seen. No excessive upconverted $1/f$ noise from the amplifier is present in the spectrum. The Q-factor of the oscillator cannot be calculated from this measurement because the minimum bandwidth of the spectrum analyzer (Agilent 4195A) is 3 Hz. The short-term stability of the oscillator was characterized using the Allen variance $\sigma(\tau)$:

$$\sigma^2(\tau) \approx \sigma^2(\tau, m) = \frac{1}{2m} \sum_{n=1}^{m} (\gamma_{[\tau(n+1)]} - \gamma_{[\tau n]})^2 \tag{1.2.44}$$

where m is the number of measurements, τ is the measurement interval, and the relative difference between two subsequent measurements γ_n is given by

$$\gamma_{[\tau n]} = \frac{f_{[\tau(n+1)]} - f_{[\tau n]}}{f_{[\tau n]}} \tag{1.2.45}$$

Using a measurement interval of 1 s and a dynamic heating power of 10 mW, the Allen variance was determined to be 9×10^{-8}. This corresponds to a short-term frequency stability of 0.04 Hz. The best values of the Allen variance reported for quartz oscillators are in the region of 10^{-11} [62].

Figure 1.2.56 shows the short-term frequency stability as a function of the dynamic heating power applied to the heating resistor. A minimum heating power of 4 mW is needed in order to achieve stable oscillation. On the other hand, increasing static heating power increases the temperature of the cantilever. The partition coefficient is thus decreased and consequently the sensitivity of the gas detection is reduced (see Figure 1.2.57). Therefore, a trade-off between stability and the temperature increase on the cantilever has to be found. This trade-off determines the type of output stage that is needed after the delay line to drive the

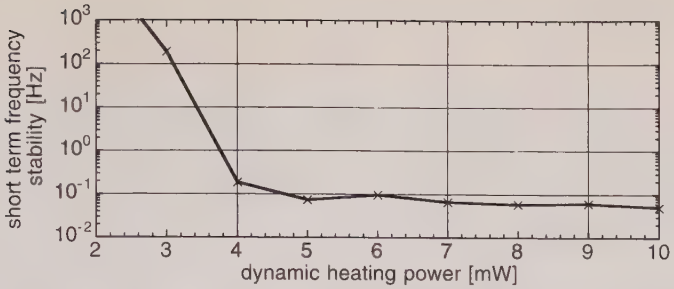


Figure 1.2.56. Short-term frequency stability of the oscillator as a function of the dynamic heating power at the resonance frequency.

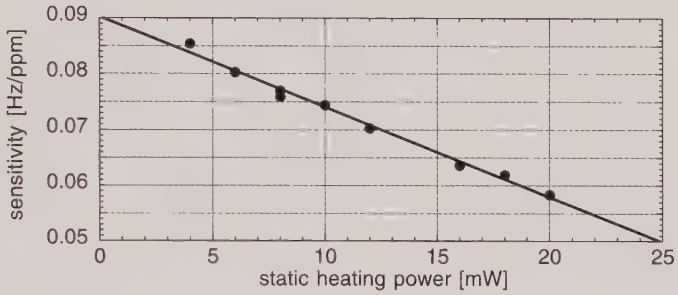


Figure 1.2.57. Sensitivity of the resonant-beam gas sensor towards octane versus static heating power.

heating resistor. Using a simple NMOS source follower with an output voltage of ~ 4 V, the nominal dynamic heating power generated by the 750 Ω heating resistor is 7.5 mW. Taking into account the process tolerances for the heating resistor and the NMOS transistor, this leads to a heating power between 6 and 9 mW.

1.2.5 Acknowledgments

The authors acknowledge D. Lange, A. Koll, N. Kerness, D. Scheiwiller, M. Zimmermann, Prof. S. Kawahito for their contributions, and the Körber Foundation, Hamburg, Germany for financial support of this project.

1.2.6 References

[1] Hagleitner, C., Hierlemann, A., Brand, O., Baltes, H., CMOS Single-chip Gas Detection Systems: Part I. In: *Sensors Update*, Vol. 11, Baltes, H., Hesse, J., Korvink, J.G. (eds.); Weinheim: Wiley-VCH, in press.

[2] Hierlemann, A., Baltes, H., Semiconductor-based Chemical Microsensors, book chapter, to be published.

[3] Hierlemann, A., Massensensitive Detektion flüchtiger organischer Substanzen mit modifizierten Polysiloxanen. *PhD Thesis*, University of Tübingen, 1996.

[4] Atkins, P.W., *Physical Chemistry*, 5th ed.; Oxford: Oxford University Press, 1994.

[5] Koll, A., CMOS Capacitive Chemical Microsystems for Volatile Organic Compounds, *PhD Thesis*, ETH Zurich, No. 13460, 1999.

[6] Misevich, K.W., *IEEE Trans. Ind. Control Instrum.* **16** (1969) 6–12.

[7] Regtien, P.P.L., *Sens. Actuators* **2** (1981) 85–95.

[8] Sheppard, N.F., Day, D.R., Lee, H.L., Senturia, S.D., *Sens. Actuators* **2** (1982) 263–274.

[9] Boltshauser, T., Baltes, H., *Sens. Actuators A* **25–27** (1991) 509–512.

[10] Kang, U., Wise, K.D., *Solid State Sensor and Actuator Workshop* (1998) 183–186.

[11] Souloff, B., *Adv. Microsyst. Autom. Appl.* **99** (1999) 262–270.

[12] Tartagni, M., Guerrieri, R., *IEEE J. Solid State Circuits* **33** (1998) 133–142.

[13] Jachowicz, R.S., Senturia, S.D., *Sens. Actuators* **2** (1981/82) 171–186.

[14] *Datasheet ADXL 105*; Analog Devices, http://www.analog.com, 1999.

[15] Sherman, S.J., Tsang, W.K., Core, T.A., Payne, R.S., Quinn, D.E., Chau, K.H.-L., Farash, J.A., Baum, S.K., *IEEE Int. Electron Device Meet.* (1992) 501–504.

[16] Lemkin, M., Boser, B.E., *IEEE J. Solid State Circuits* **34** (1999) 456–468.

[17] van der Goes, F.M.L., Meijer, G.C.M., *IEEE Trans. Instrum. Meas.* **45** (1996) 536–540.

[18] Kung, J.T., Lee, H.-S., Howe, R.T., *IEEE J. Solid State Circuits* **23** (1988) 972–977.

[19] Brandl, M., Kempe, V., *Proc. IEEE MEMS'01* (2001) 6–9.

[20] Cho, S.T., Wise, K.D., *Sens. Actuators A* **36** (1993) 47–56.

[21] Gregorian, R., Temes, G.C., *Analog MOS Integrated Circuits for Signal Processing*; New York: Wiley, 1986.

[22] Malcovati, P., CMOS Thermoelectric Sensor Interfaces, *PhD Thesis*, ETH Zurich, No. 11424, 1996.

[23] Lindell, M., Tenhunen, H., *IEEE J. VLSI Comput. Periph.* (1989) 126–128.

[24] Verhoeven, H.-J., Huijsing, J.H., *Sens. Actuators A* **52** (1996) 198–202.

[25] Kawahito, S., Maier, C., Schneider, M., Zimmermann, M., Baltes, H., *IEEE J. Solid State Circuits* **34** (1999) 1843–1851.

[26] Hagleitner, C., Koll, A., Vogt, R., Brand, O., Baltes, H., *Proc. Transducers'99* **2** (1999) 1012–1015.

[27] Suarez, E., Gray, P.R., Hodges, D.A., *IEEE J. Solid State Circuits* **10** (1975) 379–385.

[28] Wegmann, G., Vittoz, E.A., Rahali, F., *IEEE J. Solid State Circuits* **22** (1987) 1091–1097.

[29] Schiffer, B., Burstein, A., Kaiser, W.J., *IEEE J. Solid State Circuits* **33** (1998) 2134–2138.

[30] Eichenberger, C., Guggenbühl, W., *IEEE Trans. Circuits Syst.* **37** (1990) 256–264.

[31] Kawahito, S., Koll, A., Hagleitner, C., Baltes, H., Tadokoro, Y., *Trans. IEE Jpn.* **119-E** (1999) 138–142.

[32] van de Plasche, R., *Integrated Analog-to-Digital and Digital-to-Analog Converters*; Dordrecht: Kluwer, 1994.
[33] Kansy, R.J., *IEEE J. Solid State Circuits* **15** (1980) 373–375.
[34] Wey, H.M., Guggenbühl, W., *IEEE Trans. Circuits Syst.* **33** (1986) 1028–1030.
[35] Laker, K.R., Sansen, W., *Design of Analog Integrated Circuits and Systems*; New York: McGraw-Hill, 1994.
[36] Oppenheim, R.V., Schaefer, R.W., *Discrete-Time Signal Processing*; Englewood Cliffs: Prentice-Hall, 1989.
[37] Schneeberger, K., CMOS Microsystems for Thermal Presence Detection, *PhD Thesis*, ETH Zurich, No. 12675, 1998.
[38] Kerness, N., Koll, A., Schaufelbühl, A., Hagleitner, C., Hierlemann, A., Brand, O., Baltes, H., *Proc. IEEE MEMS 2000* (2000) 96–101.
[39] Kerness, N., CMOS-Based Calorimetric Chemical Microsensors, *PhD Thesis*, ETH Zurich, No. 14839, 2002.
[40] Schaufelbühl, A., Thermal Imagers in CMOS Technology, *PhD Thesis*, ETH Zurich, No. 14484, 2002.
[41] Bakker, A., High Accuracy CMOS Smart Temperature Sensors, *PhD Thesis*, TU Delft, 2000.
[42] Menolfi, C., Low Noise CMOS Chopper Instrumentation Amplifiers for Thermoelectric Microsensors, *PhD Thesis*, ETH Zurich, No. 13583, 2000.
[43] Silva-Martinez, J., Steyaert, M., Sansen, W., *High-Performance CMOS Continuous-Time Filters*; Dordrecht: Kluwer, 1993.
[44] Krummenacher, F., Joehl, N., *IEEE J. Solid State Circuits* **23** (1988) 750–758.
[45] Khorramabadi, H., Gray, P.R., *IEEE J. Solid State Circuits* **19** (1984) 939–948.
[46] Allen, P.E., Holberg, D.R., *CMOS Analog Circuit Design*; Oxford: Oxford University Press, 1987.
[47] Norshworthy, S.R., Schreier, R., Temes, G.C., *Delta Sigma Data Converters*; Piscataway, NJ, USA: IEEE Press, 1997.
[48] Chu, S., Burrus, C.S., *IEEE Trans. Circuits Syst.* **31** (1984) 913–924.
[49] Brodersen, R.W., Moscowitz, H.S., in: *VLSI Signal Processing III*; Brodersen, R.W., Moscowitz, H.S. (eds), New York: IEEE Press, 1988, pp. 523–534.
[50] Saramäki, T., Tenhunen, H., *Proc. IEEE ISCAS'88* (1988) 1525–1528.
[51] Candy, J.C., *IEEE Trans. Commun.* **34** (1986) 72–76.
[52] Antoniou, A., *Digital Filters*; New York: McGraw-Hill, 1993.
[53] *Matlab Version 5.30 for Unix*; The Math Works, 1999.
[54] Lange, D., Hagleitner, C., Hierlemann, A., Brand, O., Baltes, H., *Anal. Chem.* **74** (2002) 3084–3095.
[55] Lange, D., Brand, O., Baltes, H., *CMOS Cantilever Sensor Systems*; Berlin: Springer, 2002.
[56] Tortonese, M., Yamada, H., Barett, R.C., Quate, C.F., *Proc. Transducers'91* (1991) 448–451.
[57] Koller, S., Resonating Viscosity Sensors: Micro Versus Macro Approach, *PhD Thesis*, ETH Zurich, No. 13208, 1999.
[58] Hornung, M., Brand, O., Baltes, H., *Proc. Transducers'97* (1997) 441–444.
[59] Ohlckers, P., Holm, R., Jakobsen, H., Kvisteroy, T., Kittilsland, G., Larsen, A., Nese, M., Nilsen, S.M., Ferber, A., *Sens. Actuators A* **66** (1998) 99–104.
[60] Huijsing, J.H., *IEEE Trans. Instrum. Meas.* **25** (1976) 227–231.
[61] Huang, S.C., Ismail, M., *Proc. 37th Midwest Symp. Circuits Syst.* **2** (1994) 780–781.
[62] Lange, D., Hagleitner, C., Brand, O., Baltes, H., *Proc. Transducers'99* (1999) 1020–1023.

List of Symbols and Abbreviations

Symbol	Designation
a	activity (or effective concentration)
A	constant
Av	gain
B	signal bandwidth
c	concentration
c_A	concentration of analyte in gas phase
C_{gas}	concentration of analyte in gas phase
C_L	input capacitance
C_p	parasitic capacitance
C_{par}	sum of parasitic conductances
f_0	resonance frequency of cantilever
f_θ	output-word rate
G	Gibbs energy
ΔG^0	standard reaction Gibbs energy
gm	transconductance
h	total thickness of cantilever
H	transfer function
ΔH^0	standard enthalpy
ΔH_{vap}	enthalpy of vaporization
I_{bias}	bias current
K	equilibrium constant
K	partition coefficient
KF	flicker noise constant
l	order of filter
L	transistor length
m	number of measurements
N	number of thermocouples
n	selected input of multiplexer
n_i	stoichiometric number of component i (or amount of substance)
Q	quality factor of bandpass filter
R	resistance
S	equivalent input noise power density
S	sensitivity
S_{tn}	thermal noise power density
ΔS^0	standard entropy
ΔS_{vap}	entropy of vaporization
t_L	thickness of polymer
T	temperature
T_s	sampling time
VCM	common-mode voltage

Symbol	Designation
Vds	drain source voltage
Vd_{sat}	saturation voltage
VF_C	reactive volume fraction
$V_{polymer}$	polymer volume
W	transistor width
X	input signal
ε	dielectric constant
ε	relative mismatch between filter and chopping frequency
ϕ	phase
γ	activity coefficient
γ	Seebeck coefficient
γ_n	relative difference between two measurements
μ_i	chemical potential of component i
π	piezoresistive coefficient
ρ	density
ρ_{mean}	average-specific mass density of cantilever layer stack
σ	stress
τ	measurement interval
ω	frequency

Abbreviation	Explanation
AC	alternating current
CDS	correlated double sampling
CMFB	common-mode feedback
CMOS	complementary metal oxide semiconductor
CMRR	common-mode rejection ratio
DC	direct current
DDA	differential difference amplifier
DSP	digital signal processor
FEM	finite element method
FIR	finite-impulse response
GBW	gain-bandwidth product
MOS	metal oxide semiconductor
NMOS	n-type metal oxide semiconductor
NTF	noise power transfer function
OTA	operational transconductance amplifier
PDMS	poly(dimethylsiloxane)

Abbreviation	Explanation
PEUT	poly(etherurethane)
PMCPS	poly[methyl(cyanopropyl)siloxane]
PMOS	poly(methylsiloxane)
PMOS	p-type metal oxide semiconductor
SEM	scanning electron microscope
SNR	signal-to-noise ratio
TC	temperature coefficient
VOC	volatile organic compound

1.3 Ionomer-coated Electrodes and Nanoelectrode Ensembles as Electrochemical Environmental Sensors: Recent Advances and Prospects

Paolo Ugo, Ligia Maria Moretto and Francesca Vezzà,
Department of Physical Chemistry, University of Venice, Venice, Italy

Abstract

This article focuses on recent advances in the use of ionomer-coated electrodes and nanoelectrode ensembles for the determination of redox ions at trace concentration level with special stress on applications for environmental monitoring. After an overview on the most recent developments of ionomer-coated electrodes, attention is directed towards their use for speciation analysis in environmental samples. In particular, some examples such as the selective determination of inorganic and metallorganic species of different toxicity and the redox state speciation of multiply charged ions are discussed. The second part of the article deals with the development and electroanalytical use of a new kind of nanostructured electrode, named nanoelectrode ensembles, which are prepared by template synthesis within the pores of microporous guest membranes. Advantages, peculiarities, and limits of such devices are discussed together with perspectives for their use in electroanalytical sensing at trace and ultratrace concentration levels in environmental samples.

Keywords: coated electrodes; environmental monitoring; ion exchange; nanoelectrodes; voltammetry

Contents

1.3.1	Ion-exchange Voltammetry at Ionomer-coated Electrodes	124
1.3.2	Speciation Analysis by IEV	127
1.3.3	Molecular Engineering of the Polymer Layer	130
1.3.4	Improving Detection Limits	132
1.3.5	Ensembles and Arrays of Nanoelectrodes	132
1.3.6	Summary and Outlook	135
1.3.7	Acknowledgment	136
1.3.8	References	136

The ability to build micro- and nanostructures on electrode surfaces has given rise to new electrochemical devices in which the properties of the electrode are tailored for specific purposes [1, 2]. In principle, this is a very attractive prospective for developing electrochemical sensors suited especially for environmental analysis in complex matrices, where the presence of numerous interferences hinders the reliable use of "classical" electrodes. The vocational samples for electroanalytical methods are wet samples, that already contain a certain amount of water acting as supporting electrolyte. This does not necessarily limit their application only to saline waters (which, anyway, is a wide application field [3–5]), since most recent advances have shown that dynamic electroanalytical methods can be used for direct measurements and "in situ" determinations of redox species in a variety of unconventional matrices of environmental interest, such as rain [6–8], ice [9], lake waters [10, 11], marine sediment pore waters [12], lagoon and estuarine pore waters [13–15], hydrothermal vents [4], soils [16], and a variety of biological samples including cells and tissues [17, 18] and even frog's eggs [19, 20].

In complex natural samples many problems arise from the presence of a variety of interfering substances that, from an electroanalytical viewpoint, can be grouped as follows:

- Electroactive substances which are reduced (or oxidized) at potentials close to values typical for the analyte;
- Organics and surface active agents which can adsorb on electrode surfaces, blocking or poisoning the electrode surface;
- Ligands which complex the analyte, causing relevant shifts in its electroactivity potential.

In environmental chemistry the terms trace and ultratrace concentrations are often used; here trace is used as synonym for submicromolar ($< 10^{-6}$ M) and ultratrace for

subnanomolar ($< 10^{-9}$ M) concentrations. Note that operating at so low analyte concentration levels, if one can match with uncompensated electrical resistance effects (for instance, using ultramicroelectrodes [6, 9] or electrodes coated with polyelectrolytes [21]), low concentrations of nonelectroactive dissolved salts (included H^+ and OH^- coming from water autoprotolysis) in the millimolar range, and even slightly lower, can act as suitable supporting electrolytes.

The bioavailability and toxicity of many pollutants is related to the chemical species under which the toxic element is present in the environment, where the chemical species is a specific form of an element defined as oxidation state, complex, or molecular structure [22]. This underlines the importance that trace and ultratrace detection capabilities of environmental sensors be accompanied by speciation analysis capabilities. According to the above-cited IUPAC guidelines [22], speciation analysis is meant as the analytical activity aimed to identifying and/or measuring the quantities of one or more individual chemical species in a sample.

Voltammetric and amperometric methods of analysis have some advantageous peculiarities [23–25] which make them good candidates for environmental sensing. First, they take advantage of the immediate transduction of chemical information into an electrical signal, since the Faradaic current signal is produced directly at the electrode–sample interface. Moreover, they are sensitive and characterized by low detection limits; it is possible, in fact, to perform the preconcentration of the analyte directly on the transducer (electrode) surface via Faradaic [26] (for example, anodic (ASV) or cathodic (CSV) stripping voltammetry) and/or non-Faradaic (including adsorptive, ion exchange, complexation) preconcentration methods [27]. Moreover, they can give information on the redox state and complexation of the analyte in the sample [28]. They can be used in turbid or colored samples and do not suffer from interference from ambient light. All these characteristics make, in principle, electrochemical methods and sensors attractive for real time monitoring of trace pollutants in the environment.

Efforts need also be made to overcome the widespread use of mercury as an electrode material. Mercury has some undoubted advantages in electroanalysis, however because of its high toxicity its use should be avoided where possible. Moreover, static mercury drop electrodes are rather complicated for setting up in unmanned automatic analyses and mercury cannot be used at positive potential values because of the oxidation of mercury itself.

In order to widen the application field of electroanalytical methods it is important to develop novel electrode systems characterized by high preconcentration capabilities, good reproducibility, low cost, and the use of nontoxic materials. Ideally, they should have intrinsic molecular recognition properties so that they can give specific responses in complex real samples and can be easily used for "in field" and automatic monitoring.

The present review, which does not intend to be exhaustive, presents and discusses recent advances in the preparation, characterization and applications of polymer-coated electrodes and nanoelectrode ensembles, has the goal of trying to fulfill (at least in part) some of the above mentioned, but not easily achievable, presuppositions.

1.3.1 Ion-exchange Voltammetry at Ionomer-coated Electrodes

The first studies on ion-exchange voltammetry (IEV) started during the late 1970s/early 1980s [29–31], with the development of modified electrodes (usually glassy carbon, gold, or platinum disk electrodes) coated by a thin layer (0.1–2 μm thick) of an ion-exchange polymer which allowed the quick preconcentration and simultaneous voltammetric detection of ion redox analytes [32]. The voltammetric signal recorded at these polymer-coated electrodes depends on the concentration and apparent diffusion coefficient (D_{app}) of electroactive species incorporated in the polymeric layer. If an ion exchanger characterized by proper selectivity is used, it is possible to determine trace and ultratrace levels of ionic electroactive analytes [32]. Moreover, the presence of a polyionic coating on the electrode surface prevents adsorption or precipitation phenomena [33, 34]. In the case of analysis in complex matrices such as environmental samples, biological and organic components can, in fact, deactivate the electrode surface preventing the direct application of voltammetric methods to real world analytical challenges.

Principles and many analytical applications of IEV at polymer-coated electrodes have been reviewed a few years ago, and the readers are referred to [32] for a summary of the situation in this field up to 1995. Some of the most recent advances concerned with the use of IEV at ionomer-coated electrodes are listed in Tables 1.3.1 and 1.3.2.

From an analytical viewpoint it can be noted that the ion-exchange modifier most widely used is still Nafion. Nafion, Tosflex, and sulfonated polyesters (see Scheme 1) belong to the ionomer group, ionic polymers characterized by a small content of ionic groups [35].

As sketched in Figure 1.3.1 for the case of Eastman AQ55, these ionic groups generally interact or associate to form ion-rich hydrophilic regions surrounded by hydrophobic domains [36]. The ease of modification of the electrode surface and mechanical and chemical stability of the film of ionomers is the basis of the success of IEV with this class of coating materials.

From Tables 1.3.1 and 1.3.2 it is evident that the IEV of inorganic ions still prevails over its application to organics. It is worth emphasizing that interesting advances dealing with the extension of ion-exchange voltammetry to the electrochemical biosensors world appeared recently [57–61].

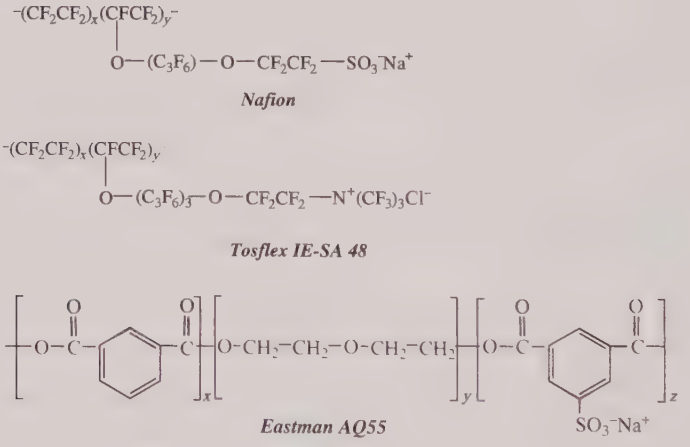
In the following we will discuss more specifically some peculiar advances in IEV concerned with the implementation of speciation capabilities and molecular engineering of the coating layer.

Table 1.3.1. Some recent examples of application of IEV to inorganic ionic analytes.

Analyte	Modifier	Comments
Pb^{II}	Nafion [37]; Nafion/1-(2-pyridylazo)-2-naphthol [38]; Quaternized poly(4-vinylpyridine) (QPVP) [39]	The chelating agent in Nafion increases sensitivity and selectivity [38]. Interference from surfactants and other metal ions is reduced [38]. Anionic complexes generated locally in the coating are preconcentrated by the QPVP; the method is applied to groundwater analyses [39].
Cu^{II}, Cd^{II}, Pb^{II}, Zn^{II}	Nafion and Nafion/poly(vinylsulphonic acid) [40]; Polyester sulphonate [41]	Comparison of different kinds of coatings for batch analysis of environmental samples such as river and run-off waters. Effect of surfactants was examined.
CH_3Hg^+, $HgCl_4^{2-}$	Nafion [42]; Tosflex [43]	DPV and multiple square wave voltammetry (MSWV) allowed ultratrace trace determination of CH_3Hg^+ with Nafion [42]. The role of Tosflex in the redox mechanism was examined. For irreversible systems, lowering of diffusion coefficients in the coating lowered reduction overpotential [43].
Tl^{I}	Nafion [44]	Square wave stripping, in the presence of added EDTA to eliminate interference from other metal cations. Detection limit (DL) is 0.01*ppb.
Eu^{III}	Nafion [45]	Optimization of the MSWV parameters for ultratrace determination. DL = 10 pm. Application to analysis in cooling waters of nuclear plants.
Bi^{III}	Poly(4-vinylpyridine) [46]; Tosflex [47]	Square wave stripping of anionic complexes of Bi^{III} with chloride. No deoxygenation and reduced interference from surfactants [46]. DL with Tosflex is 0.58 ppb [47]
Zn^{II}	Tosflex [48]	Square wave stripping with Tosflex-coated electrodes, which protects the electrode from organic interference. DL 0.1 ng mL^{-1}.
Te^{IV}	Tosflex/8-quinolinol [49]	Cathodic stripping of anion $TeCl_6^{2-}$ combination of IEV, chelation and electrocatalysis by oxine. DL 0.1 ng mL^{-1}.

Table 1.3.2. Recent examples of application of IEV to organic redox analytes.

Analyte	Modifier	Comments
Viologen	Nafion/clay [50]; Nafion [51]; Nafion/tetraethoxysilane [52]; Polyester sulphonate [53]	Good performances for the composite with clay [50] applied to groundwater and lake waters. Electrocatalytic determination by reduction of added permanganate [51]. Accumulation and release of the analyte at the composite with silane is faster than with pure Nafion. DL in the picomolar range [52]. Study of factors influencing the preconcentration and determination at modified electrodes in acetonitrile solution [53].
N-nitros-amines	Sulphopolyester (Eastek) [54]	Studied effect of the degree of sulfonation, comparison with Nafion, cellulose acetate, and butyrate. DL = 0.1 nM.
Parathion	Nafion [55]	Linear range up to 15 μM. DL = 50 nM.
Phenol	Nafion [56]	Adsorptive stripping in the presence of ethyltrimethylammonium bromide, which favors the analysis. DL = 1 nM.



Scheme 1. Structural formulas of ionomers used as electrode coatings.

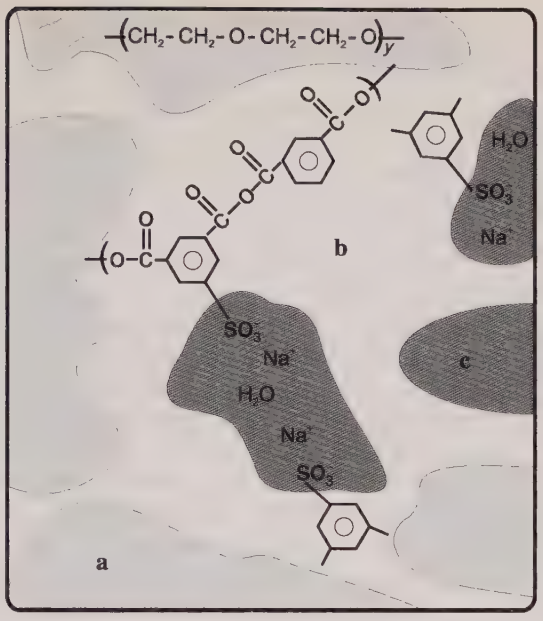


Figure 1.3.1. Sketch of the domain structure of Eastman AQ55 ionomer: (**a**) hydrophobic carbon chain domains; (**b**) interfacial transition zones; (**c**) ionic clusters.

1.3.2 Speciation Analysis by IEV

Newer examples of the contribution that IEV can offer to environmental analyses is given by its application to speciation analyses. A first attempt in this direction was made by Morrison and Florence [62], who speciated model copper complexes on the basis of size and charge differences by using mercury film electrodes coated with Nafion (cation exchanger) or cellulose acetate/Nafion (size and charge exclusion) or poly(4-vinylpyridine) (PVP, anion exchanger at pH < 5.2). In spite of the novelty of the idea, the results presented in that study were difficult to evaluate since the authors employed PVP at pH values where the pyridine groups are deprotonated and the polymer loses its anion-exchange properties [63].

Ugo et al. [64] demonstrated that charge speciation of copper–amino acid complexes can be achieved by using Nafion-coated glassy carbon electrodes to detect cationic forms and Tosflex-coated electrodes for anionic ones. Note that the ion-exchange properties of both these modifiers are independent of the solution pH.

The comparison between signals obtained at Nafion-coated and uncoated mercury film electrodes allowed Chakrabarti and co-workers to get speciation information on cadmium and lead complexation by fulvic acids in real samples, such as river waters [65, 66].

Very recently, problems related to mercury speciation in natural waters were examined. In this case, the goal is the development of methods that give responses for metallorganic species, such as methylmercury, separated from those relevant to inorganic species. This is a key point for environmental studies of this toxic metal. New examples of application of modified electrodes to trace and ultratrace mercury determinations has been published recently [67–71], however they dealt with the determination of total mercury.

Advances in methylmercury speciation issues were obtained recently with respect to the main metallorganic mercury species present in water, the monomethyl mercury cation, which is efficiently preconcentrated and detected with no Faradaic preconcentration at Nafion-coated electrodes [42]. Detection limits are in the 10^{-8} M range using differential pulse voltammetry (DPV), but can be lowered to the subnanomolar range by performing the scan at the coated electrode by a new electroanalytical technique named double differential multiple square-wave voltammetry (MSW-DD) [42]. As shown in Figure 1.3.2, at Cl^- concentrations typical of seawaters (about 0.5 M), the prevailing metallorganic mercury species is neutral CH_3HgCl while the prevailing inorganic species is the anion $HgCl_4^{2-}$ [68]. Note that the latter is preconcentrated by Tosflex-coated electrodes

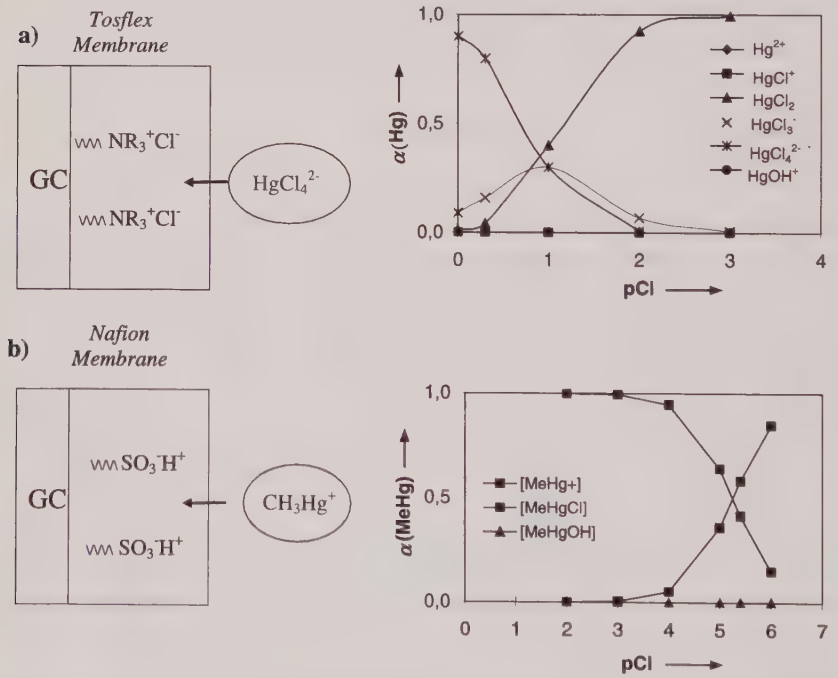


Figure 1.3.2. Species distribution diagrams and preconcentration schemes of inorganic and metallorganic mercury in acidic aqueous solutions (pH 2); preconcentration sketches refer to (**a**) CH_3Hg^+ in the absence of added NaCl, and (**b**) inorganic Hg^{II} in 0.5 M NaCl solution.

with a selectivity (given as the ratio of the sensitivity for CH_3HgCl over sensitivity for $HgCl_4^{2-}$) equal to 0.045 [72]. Overall, these results indicate that Nafion-coated electrodes can be used for the determination of CH_3Hg^+ in fresh and low salinity waters, while Tosflex-coated electrodes can be used to determine inorganic mercury in the same kind of sample after the addition of the proper amount of chloride [43] (or directly in seawater as well).

Another speciation-related issue involves the possibility of exploiting polymer-coated electrodes to perform redox-state speciation analyses. When the concentration of redox ions are relatively high (10^{-4} M), such analyses are carried out using metal indicator electrodes (typically platinum electrodes) and measuring potentiometrically what is referred to as the E_H value (see for example [73]). However, such analyses are seriously limited by the lack of sensitivity of the method and because the platinum electrode is easily poisoned by surfactants and organics present in natural samples. An alternative is offered by substituting classical E_H electrodes with polymer-coated electrodes similar to those used in IEV. Somehow, this idea retraces previous studies, in which the so called "coated wire" indicator electrodes were employed [74]. However, more information can be obtained by integrating the information coming from the same electrode, used as both a voltammetric and potentiometric sensor [75]. This makes the total quantification of signals easier and more reliable thanks to the higher sensitivity of voltammetry with respect to potentiometry, leaving to this last technique only the task of quantifying the ratio between the different redox species. Such concepts have been developed and experimented for Fe^{3+} and Fe^{2+} speciation, which is an issue of great concern in environmental chemistry, since it allows one to gain important information on the redox chemistry of anoxic waters, pore waters, and sediments as well as on the precipitation or release of toxic metals incorporated in Fe^{II} or Fe^{III} precipitates [76]. When in solution, the cations Fe^{2+} and Fe^{3+} are both preconcentrated efficiently by Nafion-coated electrodes so that they are detected by IEV down to nanomolar levels [75]. The problem here is to distinguish between the two different oxidation states. Scholtz and Hermes recently proposed a redox speciation based on the analysis of the shape of the cyclic voltammograms recorded at the electrode–solution interface when both redox partners are present in solution [77, 78]. However, the attempt to extend this methodology to the case of polymer-coated electrodes gave less clear results [79], probably because of the complexity of charge transport pathways for reactants incorporated in ionomeric coatings which can undergo redox cross-reactions, self exchange, or be characterized by disparate and/or concentration-dependent diffusion coefficients [1, 32]. Heineman et al. [80] faced a similar problem and proposed a method for obtaining information about the redox state by examining the anodic and cathodic branches of steady-state currents recorded at polymer-coated ultramicroelectrodes. However, limitations come from the requirement of knowing the apparent diffusion coefficient in the coating of both redox partners.

It appears more practical to get redox state information from a nondynamic method, such as open circuit potential measurements at the coated electrode, which under equilibrium conditions are not influenced by charge and mass transport phenomena [76, 79].

Some years ago it was shown that in evaluating these data one has to take into account the permselectivity of the Nafion coating [81, 82]. Under such condition, the redox reaction for the $Fe^{2+/3+}$ couple at the NCE can be written as Equation (1.3.1), where X^+ is the supporting electrolyte cation.

$$(R\text{-}SO_3^-)_3Fe^{3+} + X^+ + e^- \rightleftharpoons (R\text{-}SO_3^-)_2Fe^{2+} + (R\text{-}SO_3^-)X^+ \qquad (1.3.1)$$

The open circuit potential measured at the coated electrode is expected to obey Equation (1.3.2), where $E^{0\prime}$ is the formal potential of the redox couple incorporated in the coating and the subscripts p and s refer to concentrations in the polymer and solution phases, respectively.

$$E(\mathrm{OC})_{\mathrm{IEV}} = E^{0\prime} + (RT/nF)\log\{[Fe^{3+}]_p[X^+]_s/[Fe^{2+}]_p[X^+]_p\} \qquad (1.3.2)$$

Proper substitution and approximations [79] give Equation (1.3.3), where K is given by Equation (1.3.4) and $k_{\mathrm{DFe^{III}}}$ and $k_{\mathrm{DFe^{II}}}$ are relevant ion-exchange distribution coefficients which can be obtained by proper voltammetric measurements [32].

$$E(\mathrm{OC})_{\mathrm{IEV}} = K + (RT/nF)\log([Fe^{3+}]_s/[Fe^{2+}]_s) \qquad (1.3.3)$$

$$K = E^{0\prime} + (RT/nF)\log(k_{\mathrm{DFe^{III}}}/k_{\mathrm{DFe^{II}}}) + (RT/nF)\log([X^+]_s/[X^+]_p) \qquad (1.3.4)$$

Experimental results showed that open circuit potentials effectively obey Equation (1.3.3), which can be extended also to the case of redox anions incorporated within an anion exchanger such as Tosflex [79]. The method has been applied to redox speciation studies in the pore waters and sediments of the Venice Lagoon [75, 76, 79].

1.3.3 Molecular Engineering of the Polymer Layer

Advances in the reliable use of polymer-coated electrodes for analytical purposes depend on improvements in the control and reproducibility of the deposition of the polymer layer. This can also allow the lowering of the film thickness, thereby shortening response times. Advances in this direction can come from the electrostatic spray deposition method proposed by Hoyer et al. [83] that allowed the deposition of very uniform thin films (about 0.3 μm), keeping unaltered the peculiar features of Nafion as far as permselectivity and rejection of interferents are concerned.

Other possibilities are opened by the application of molecularly engineered deposition methods such as those based on Langmuir-Blodgett (LB) films [84, 85].

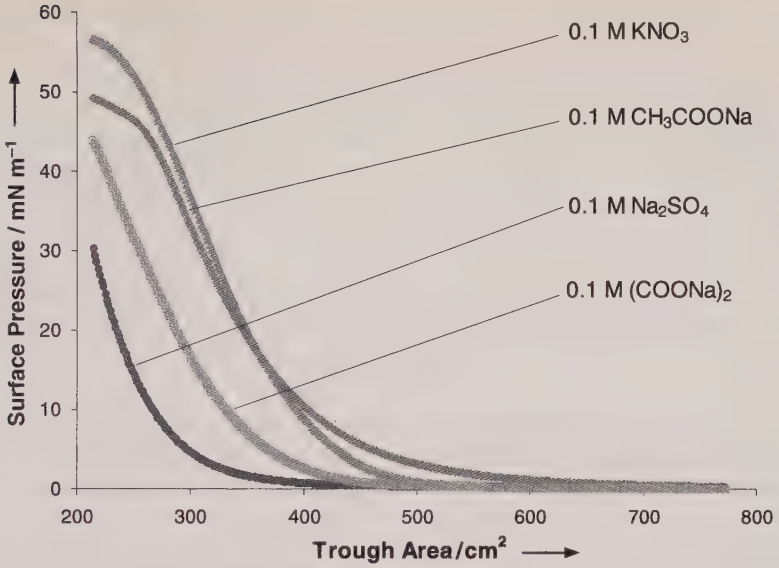


Figure 1.3.3. Surface pressure versus trough area isotherms of Tosflex in the subphases indicated. The trough area (x-axis) during the compression is directly proportional to the area per molecule [84, 85]; the smaller this parameter, the more condensed is the interfacial film. Note that for polymeric materials with distributed molecular weights it is difficult to define precisely the area per molecule values. The initial total area of the trough was 772.5 cm^2; 200 μL of 0.5 mg mL^{-1} Tosflex solution in water/isopropanol/methanol (1/1/1) were added to 1250 mL of the subphase [87]. The interfacial layer was compressed at 3 cm min^{-1}.

It was recently shown that if the proper electrolyte is dissolved in an aqueous subphase, Nafion forms an interfacial film at the air–water interface, which can be compressed in a typical Langmuir-Blodgett experiment, giving a stable film transferable to the surface of a variety of materials including electrode surfaces [86]. The thickness of the Nafion film transferred can be as thin as few nanometers [86] and scales with the number of layers deposited. The film keeps the performances typical of thicker Nafion coatings as far as permselectivity and incorporation of multiply charged cations are concerned. However because of the much smaller thickness of the coating, ion-exchange equilibration is faster.

Langmuir-Blodgett and Langmuir-Schaefer [84, 85] deposition procedures can be used also for other ionomers such as Tosflex and Eastman AQ55 [87]. The condensation degree of the interfacial film depends on the nature of the electrolyte dissolved in the subphase. The Langmuir isotherms reported in Figure 1.3.3 and obtained for Tosflex [87] show that the condensation degree depends on the anion dissolved in the subphase according to the sequence $SO_4^{2-} \gg (COO^-)_2 \gg NO_3^- > CH_3COO^-$.

This sequence parallels the sequence of the ion-exchange selectivity of strong anion exchangers [88], so suggesting that the same factor (namely charge to ra-

dius ratio) rules both phenomena. Note that interfacial films with higher condensation degrees give more stable ultrathin coatings when transferred to electrode surfaces [86, 87].

The Langmuir-Blodgett technique looks like an attractive method to control the deposition of ionomers which can be employed to prepare large numbers of modified electrodes, starting, for instance, from plates containing multiply screen-printed electrodes [89, 90] all modified in just one LB experiment.

1.3.4 Improving Detection Limits

Ultratrace detection capabilities typical of IEV can be further improved using electroanalytical techniques and devices characterized by higher signal-to-noise ratio with respect to conventional ones. For brevity, in this review we will only mention the use of multiple square-wave voltammetry (MSWV) [91–93], whose implementation in the double differential mode (MSW-DD) [42] has been shown to be useful for further lowering detection limits achievable with polymer-coated electrodes [45, 75].

A different, but complementary, approach for reaching the same goal is offered by the use of electrode systems containing metallic components of nanoscopic dimensions such as those presented below.

1.3.5 Ensembles and Arrays of Nanoelectrodes

Nanoelectrode ensembles (NEEs) and arrays (NEAs) are new electroanalytical tools that can find application in a variety of fields ranging from sensors to electronics, from energy storage to magnetic materials [2]. The characteristics that distinguish NEEs and NEAs from conventional macro (millimeter-sized) or even ultramicro (micrometer-sized) electrodes are:

- Dramatic lowering of capacitive currents which act as noise in voltammetry [94, 95].
- Extreme sensitivity to the kinetics of the charge transfer process [96], which means capability to measure very high charge transfer rate constants [97].

In case of experimental inability to get direct morphological information on the NEE structure by electron or scanning probe microscopy [98], the lack of some of these characteristics, such as the persistence of high capacitive currents, should be taken as a diagnostic indication of a failure in the preparation procedure.

The difference between NEEs and NEAs is that in the former the spatial distribution of the nanoelectrodes is random while in the array it is ordered. In the following we will refer generically to NEEs (which are easier to prepare), with the majority of the considerations being transferable to NEAs too. There are two typical methods which proved to be successful in preparing NEEs with the above characteristics. In the "template synthesis" of NEEs developed by Martin et al. [2], gold is deposited into the pores of nanoporous membranes, typically by electroless plating procedures. Historically, the template synthesis was introduced by Possin [99] and refined by Williams and Giordano [100], who prepared different metallic wires with diameters as small as 10 nm within the pores of etched nuclear damage tracks in mica. Electrochemical deposition of metal (nickel and cobalt) nanowires [101] was performed as well as nanofibers of semiconductors [102], oxides [103], and conducting polymers [2].

Other approaches are based on exploiting the defects generated in self-assembled monolayers as NEEs [97, 104, 105]. NEAs have also been fabricated by creating and controlling the pores in a block copolymer self-assembled matrix [106]. As summarized in Figure 1.3.4, NEEs can exhibit three distinct voltammetric response regimes depending on the scan rate or reciprocal distance between the nanoelectrode elements [107]. When radial diffusion boundary layers overlap totally (radius of diffusion hemisphere larger than average hemidistance between electrodes) NEEs behave as planar macroelectrodes with respect to Faradaic currents. When diffusion hemispheres become shorter (higher scan rates), the current response is dominated by radial diffusion at each single element. At very high scan rates, the linear active state is reached in which the current response is governed by linear diffusion to the individual nanodisk. Transition from one regime to the other as a function of nanoelements distance was demonstrated experimentally [108].

When trying to develop polymer-coated NEEs [95], as far as detection limits are concerned the most advantageous regime is the total overlap regime, since this regime gives the highest signal to noise ratio. If we consider a linear sweep voltammetric experiment involving a reversible redox couple, peak currents (the signal, S) and capacitive currents (the main component of the noise, N) at a NEE operating under a total overlap regime are given in Equations (1.3.5) and (1.3.6), where A_{act} is the active area (nanodisks's surface) and A_{geom} is the total geometric area of the ensemble (nanodisks and insulator).

$$S = I_p = 2.69 \times 10^5 n^{3/2} A_{geom} \nu^{1/2} D^{1/2} C_b \tag{1.3.5}$$

$$N = I_c = A_{act} \nu C_d \tag{1.3.6}$$

At a conventional electrode, in the same experimental conditions, Equations (1.3.7) and (1.3.8) hold.

$$S = I_p = 2.69 \times 10^5 n^{3/2} A_{geom} \nu^{1/2} D^{1/2} C_b \tag{1.3.7}$$

$$N = I_c = A_{geom} \nu C_d \tag{1.3.8}$$

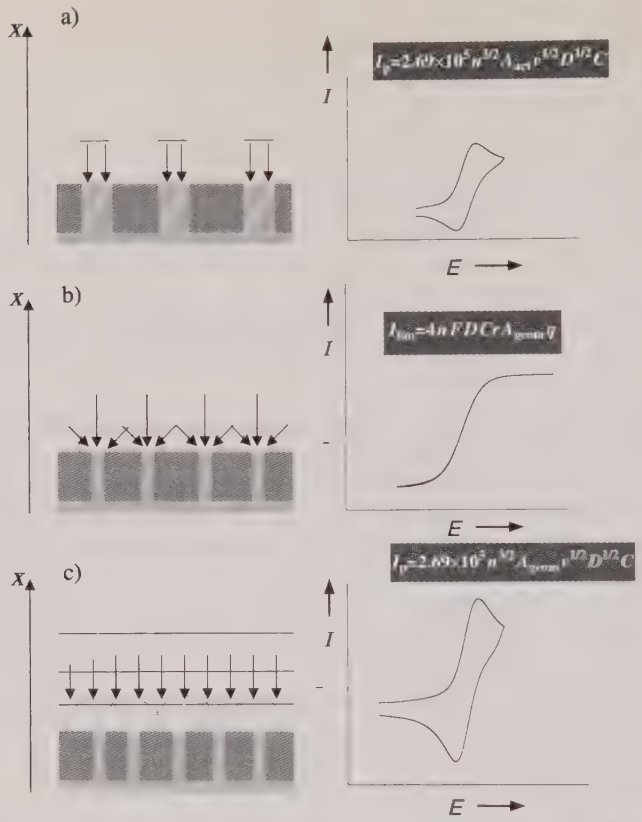


Figure 1.3.4. Schematic drawing of the different diffusional regimes at NEE: (**a**) linear active; (**b**) pure radial; (**c**) total overlap. The time scale (scan rate) of the experiment or the distance between nanodisk elements lower from cases (**a**) to (**c**). Relevant equations for peak currents (cases (**a**) and (**c**)) and plateau current (case (**b**)) refer to reversible redox systems. A_{act} is the active area (nanodisks's surface), A_{geom} is the total geometric area of the ensemble (nanodisks and insulator), q is the nanodisk density (disk cm^{-2}), and all other symbols have their usual meaning.

Combining Equations (1.3.5)–(1.3.8) gives Equation (1.3.9).

$$(S/N)_{NEE} = (S/N)_{conv} A_{geom}/A_{act} \tag{1.3.9}$$

Equation (1.3.9) gives evidence that the signal-to-noise ratio at the NEE is higher than that at a conventional electrode of the same geometric area with a proportionality factor that is the reciprocal of the fractional electrode area $f = A_{act}/A_{geom}$. For NEEs prepared with the template method, f values are generally between 1×10^{-3} and 5×10^{-3} [94, 95]. This means that detection limits with such devices are two to three orders of magnitude lower than with conventional electrodes [95].

For a fixed A_{geom}, S is maximum when a total overlap regime is operative, being lower in the case of a pure radial regime. In this case, in fact, only a certain percentage of the geometric area of the ensemble contributes to producing a Faradaic current, while in the total overlap regime this percentage is 100%. Note that Faradaic currents at NEEs in the total overlap regime are identical to those at conventional electrodes of the same geometric area as the ensemble, and the improvement produced by the use of the NEE is all in the dramatic lowering of the capacitive current [94].

From a detection limit viewpoint and with the perspective of using polymer-coated NEEs [95], a second relevant advantage comes from the lower weight of the diffusion coefficient in determining the response under linear diffusion (as in the total overlap case) with respect to radial diffusion control. In the total overlap regime, responses obey the Randles-Sevcik equation [96] with peak currents depending on $D^{1/2}$. Note that for coated electrodes D is really the apparent diffusion coefficient D_{app} [32], at polymer-coated electrodes. Under radial conditions, signals depend directly on D_{app} [109].

The amplification of signals typical of IEV is related to the preconcentration capability of the modified electrode, but it is somehow contrasted by the lowering of D_{app} in the polymer with respect to D in solution [32]. It is therefore apparent that the diffusional regime to be preferred is the total overlap regime, since this is the regime displaying the lowest dependence of signals on D_{app} ($D_{app}^{1/2}$ for total overlap versus D_{app} for pure radial diffusion) [109].

The ability of NEEs to furnish well-resolved cyclic voltammograms for trace redox species has interesting consequences exceeding detection capabilities. This was shown in the case of redox species which undergo unwanted adsorption on electrode surfaces [110]. If adsorption is concentration dependent (for example, adsorption increases with the solution concentration of the electroactive molecule), then lowering the solution concentration below the adsorption limit can solve the problem. This was shown to be the case for some phenothiazines used as redox mediators in biosensors [110]. At the macro electrode, the voltammograms for the phenothiazines are distorted by adsorption of the reduction products. This adsorption process is concentration dependent. The lower detection limits obtained at the NEE (for example 0.12 μm for azure A by direct voltammetry with no preconcentration) allowed for the use of lower mediator concentrations and, as a result, this unwanted adsorption process was eliminated.

1.3.6 Summary and Outlook

Ion-exchange voltammetry at ionomer-coated electrodes is progressively growing as a widespread powerful electroanalytical technique. It is suitable for trace and ultratrace environmental monitoring and can be used as a valuable tool in speciation analysis. The ionomer coating gives the electrode high sensitivity and selec-

tivity and protects it from surfactants and organics interferences. Further efforts should be devoted in the near future to improving the reproducibility and control of the coating deposition procedure and to shorten the time lag required for the ion-exchange preconcentration step. The application of molecular engineering procedures, such as LB, electrostatic spray, and layer by layer deposition, can play a crucial role in this issue.

Nanoelectrode ensembles show dramatically enhanced signal-to-noise ratios with respect to any other electrode system. The combination of ion-exchange voltammetry and NEEs appears to be a powerful tool able to push voltammetric techniques towards simpler, faster, and more sensitive and selective detection schemes.

Really, electroanalysis with NEEs is still in its infancy, however, since the potential of such devices is great, we believe that with the development of simpler methods of preparation their application will widen quickly.

1.3.7 Acknowledgment

Partial financial support from MURST (Rome) is acknowledged. Collaboration with Chuck Martin (University of Florida, Gainesville) and Jean Chevalet (Université Pierre et Marie Curie, Paris) is a pleasure to acknowledge.

1.3.8 References

[1] Murray, R., *Molecular Design of Electrode Surfaces*, New York: Wiley, 1992.
[2] Martin, C.R., in *Electroanalytical Chemistry, vol.21* (Bard, A.J., Rubinstein, I., eds.), New York: Marcel Dekker, 1999, pp.1–74.
[3] Van den Berg, C.M.G. in *Chemical Oceanography, vol.9*, Riley, J.P. (ed.), London: Academic Press, 1989, pp.196–245.
[4] Taillefert, M., Luther III, G.W., Nuzzio, D.B., *Electroanalysis* **12** (2000) 401–412.
[5] Tercier, M.L. Buffle, J., *Electroanalysis* **5** (1993) 187–200.
[6] Daniele, S., Baldo, M.A., Ugo, P., Mazzocchin, G.A., *Anal. Chim. Acta* **219** (1989), 19–26.
[7] Daniele, S., Bragato, C., Baldo, M.A., *Anal. Chim. Acta* **346** (1997) 145–156.
[8] Emons, H., Baade, A., Schoning, M.J., *Electroanalysis* **12** (2000) 1171–1176.
[9] Bond, A.M., Pfund, V.B., *J. Electroanal. Chem.* **335** (1992) 281–295.
[10] Tercier, M.-L., Buffle, J., Graziottin, F., *Electroanalysis* **10** (1998) 355–363.
[11] Tercier-Waeber, M.-L., Belmont-Hebert, C., Buffle, J., *Environ. Sci. Technol.* **32** (1998) 1515–1521.
[12] Luther III, G.W., Reimers, C.E., Nuzzio, D.B., Lovalvo, D., *Environ. Sci. Technol.* **33** (1999) 4352–4356.
[13] Brendel, P.J., Luther III, G.W., *Environ. Sci. Technol.* **29** (1995) 751– 761.

[14] Bertolin, A., Mazzocchin, G.A., Rudello, D., Ugo, P., *Mar. Chem.* **59** (1997) 127–140.
[15] Ugo, P., Bertolin, A., Moretto, L.M., *Int. J. Environ. Anal. Chem.* **73** (1999) 129–143.
[16] Svegl, I.G., Ogorevc, B., *Fresenius J. Anal. Chem.* **367** (2000) 701–706.
[17] Amatore, C., Kelly, R.S., Kristense, E.W., Kuhr, W.G., Wightman, R.M., *J. Electroanal. Chem.* **213** (1986) 31–42.
[18] Fabre, B., Burlet, S., Cespuglio, R., Bidan, G., *J. Electroanal. Chem.* **426** (1997) 75–83.
[19] Fung, Y.S., Wong, C.C.W., Deng, J.Q., *Electrochim. Acta* **42** (1997) 3257–3264.
[20] Fung, Y.S., Mak, J.L.L., *Electrochim. Acta* **44** (1999) 3855–3864.
[21] Bontempelli, G., Comisso, N., Toniolo, R., Schiavon, G., *Electroanalysis* **9** (1997) 433–443.
[22] Templeton, D.M., Ariese, F., Cornelis, R., Danielsson, L.-G., Muntau, H., Van Leeuwen, H.P., Lobinski, R., *Pure Appl. Chem.* **72** (2000) 1453–1470.
[23] Bard, A.J., Faulkner, L., *Electrochemical Methods, 2nd ed.*, New York: Wiley, 2000.
[24] Wang, J., *Analytical Electrochemistry, 2nd ed.*, Weinheim: Wiley-VCH, 2000.
[25] Kissinger, P.T., Heineman, W.R., *Laboratory Techniques in Electroanalytical Chemistry, 2nd ed.*, New York: Marcel Dekker, 1999.
[26] Brainina, Kh., Neyman, E., *Electroanalytical Stripping Methods* in *Chemical Analysis*, New York: Wiley, 1993.
[27] Wang, J. in *Electroanalytical Chemistry, vol. 16*, Bard, A.J. (ed.), New York: Marcel Dekker, 1989, pp. 1–88.
[28] Buffle, J., *Complexation Reactions in Aquatic Systems – An Analytical Approach*, Chichester: Ellis Horwood, 1988).
[29] Oyama, N., Anson, F.C., *J. Electrochem. Soc.* **127** (1980) 247–250.
[30] Rubinstein, I., Bard, A.J., *J. Am. Chem. Soc.* **102** (1980) 6641–6642.
[31] Espencheid, M.W., Ghatak-Roy, A.R., Moore III, R.B., Penner, R.M., Szentirmay, M.N., Martin, C.R., *J. Chem. Soc. Faraday Trans. I* **82** (1986) 1051–1070.
[32] Ugo, P., Moretto, L.M., *Electroanalysis* **7** (1995) 1105–1113.
[33] Hoyer, B., Florence, T.M., Batley, G.E., *Anal. Chem.* **59** (1987) 1608–1614.
[34] Hoyer, B., Florence, T.M., *Anal. Chem.* **59** (1987) 2839–2842.
[35] Kroschwitz, J.I., *Concise Encyclopedia of Polymer Science and Engineering*, New York: Wiley, 1990, pp. 495–497.
[36] Yeo, R.S., Yeager, H.L., *Modern Aspects of Electrochemistry, Vol. 16* (Conway, B.E., White, R.E., Bockris J.O'M., eds.), New York: Plenum Press, 1985, pp. 437–504
[37] Zen, J.-M., Ting, Y.-S., *Anal. Chim. Acta* **332** (1996) 59–65.
[38] Hu, Z., Seliskar, C.J., Heineman, W.R., *Anal. Chim. Acta* **369** (1998) 93–101.
[39] Zen, J.-M., Chung, H.-G., Ilagovan, G., *Electroanalysis* **11** (1999) 108–113.
[40] Brett, C.M.A., Fungaro, D.A., Morgado, J.M., Gil, M.H., *J. Electroanal. Chem.* **468** (1999) 26–33.
[41] Brett, C.M.A., Fungaro, D.A., *Talanta* **50** (2000) 1223–1231.
[42] Moretto, L.M., Ugo, P., Lacasse, R., Champagne, G.Y., Chevalet, J., *J. Electroanal. Chem.* **467** (1999) 193–202.
[43] Moretto, L.M., Mazzocchin, G.A., Ugo, P., *J. Electroanal. Chem.* **427** (1997) 113–121.
[44] Lu, T.-H., Yang, H.-Y., Sun, I.-W., *Talanta* **49** (1999) 59–68.
[45] Moretto, L.M., Chevalet, J., Mazzocchin, G.A., Ugo, P., *J. Electroanal. Chem.* **498** (2001) 117–126.

[46] Zen, J.-M., Chung, M.-J., *Anal. Chim. Acta* **320** (1996) 43–51.
[47] Yang, H.-Y., Chen, W.-Y., Sun, I.-W., *Talanta* **50** (1999) 977–984.
[48] Lu, T.-H., Huang, J.-F., Sun, I.-W., *Anal. Chim. Acta* **454** (2001) 93–100.
[49] Lu, T.-H., Sun, I.-W., *Electroanalysis* **10** (1998) 1052–1056.
[50] Zen, J.M., Jeng, S.H., Chen, H.J., *Anal. Chem.* **68** (1996) 498–502.
[51] Lu, T.-H., Sun, I.-W., *Talanta* **53** (2000) 443–451.
[52] Barroso-Fernandez, B., Lee-Alvarez, M.T., Seliskar, C.J., Heineman, W.R., *Anal. Chim. Acta* **370** (1998) 221–230.
[53] Brunetti, B., Ugo, P., *J. Electroanal. Chem.* **460** (1999) 38–45.
[54] Collyer, S.D., Bradbury, S.E., Hatfield, J.V., Higson, S.P.J., *Electroanalysis* **13** (2001) 332–337.
[55] Zen, J.-M., Jou, J.-J., Kumar, A.S., *Anal. Chim. Acta* **396** (1999) 39–44.
[56] Yi, H., Wu, K., Hu, S., Cui, D., *Talanta* **55** (2001) 1205–1210.
[57] Glazier, S.A., Campbell, E.R., Campbell, W.M., *Anal. Chem.* **70** (1998) 1511–1515.
[58] Ugo, P., Zangrando, V., Moretto, L.M., Brunetti, B., *Biosens. Bioelectron.* **17** (2002) 479–482.
[59] Lojou, E., Luciano, P., Nitsche, S., Bianco, P., *Electrochim. Acta* **44** (1999) 3341–3352.
[60] Authier, L., Schollhorn, B., Moiroux, J., Limoges, B., *J. Electroanal. Chem.* **488** (2000) 48–58.
[61] Haiying, L., Tailin, Y., Kang, S., Haihong, L., Deyao, Q., *Anal. Chim. Acta* **344** (1997) 187–199.
[62] Morrison, G.M.P., Florence, T.M., *Electroanalysis* **1** (1989) 485–491.
[63] Zumbrunnen, H.R., Anson, F.C., *J. Electroanal. Chem.* **152** (1983) 111–124.
[64] Ugo, P., Moretto, L.M., Mazzocchin, G.A., *Anal. Chim. Acta* **273** (1993) 229–236.
[65] Lam, M.T., Chakrabarti, C.L., Cheng, J., Pavski, V., *Electroanalysis* **9** (1997) 1018–1029.
[66] Murimboh, J., Lam, M.T., Hassan, N.M., Chakrabarti, C.L., *Anal. Chim. Acta* **423** (2000) 115–126.
[67] Walcarius, A., Bessierre, J., *Chem. Mater.* **11** (1999) 3009–3011.
[68] Ugo, P., Moretto, L.M., Bertoncello, P., Wang, J., *Electroanalysis* **105** (1997) 1017–1021.
[69] Kula, P., Navratilova, Z., Kulova, P., Kotoucek, M., *Anal. Chim. Acta* **385** (1999) 91–101.
[70] Stojko, N.-Yu., Brainina, Kh.Z., Faller, C., Henze, G., *Anal. Chim. Acta* **371** (1998) 145–153.
[71] Turian, I., Atiya, M., Mandler, D., *Electroanalysis* **13** (2001) 653–659.
[72] Moretto, L.M., Ugo, P., unpublished data.
[73] Teasdale, P.R., Minett, A.J., Dixon, K., Lewis, T.W., Batley, G.E., *Anal. Chim. Acta* **367** (1998) 201–213.
[74] Cunningham, L., Freiser, H., *Anal. Chim. Acta* **180** (1986) 271–279.
[75] Ugo, P., Moretto, L.M., Rudello, D., Birriel, E., Chevalet, J., *Electroanalysis* **13** (2001) 661–668.
[76] Ugo, P., Cavalieri, F., Rudello, D., Moretto, L.M., Argese, E., *Sensors* **1** (2001) 102–113.
[77] Scholz, F., Hermes, M., *Electrochem. Commun.* **1** (1999) 345–348.
[78] Scholz, F., Hermes, M., *Electrochem. Commun.* **2** (2000) 814.
[79] Ugo, P., Moretto, L.M., De Boni, A., Scopece, P., Mazzocchin, G.A., *Anal. Chim. Acta*, in press.
[80] Hu, Z., Heineman, W.R., *Anal. Chem.* **72** (2000) 2395–2400.

[81] Naegeli, R., Redepenning, J., Anson, F.C., *J. Phys. Chem.* **90** (1986) 6227–6232.
[82] Redepenning, J., Anson, F.C., *J. Phys. Chem.* **91** (1987) 4549–4553.
[83] Hoyer, B., Jensen, N., Busch, L.P., *Electroanalysis* **13** (2001) 843–848.
[84] Petty, M.C., *Langmuir-Blodgett Films*, Cambridge: University Press, 1996.
[85] Ulman, A., *Ultrathin Organic Films*, San Diego: Academic Press, 1991.
[86] Bertoncello,P., Ram, M.K., Notargiacomo, A., Ugo, P., Nicolini, C., *Phys. Chem. Chem. Phys.* **4** (2002) 4036–4043.
[87] Vezzà, F., Thesis, Università Cà Foscari di Venezia (Italy), 2002).
[88] Helfferich, F., *Ion Exchange*, New York: McGraw-Hill, 1962, p.168
[89] Wring, S.A., Hart, J.P., *Analyst (Cambridge, U.K.)* **117** (1992) 1281–1286.
[90] Wang, J., Lu, J., Tian, B., Yarnitzky, C., *J. Electroanal. Chem.* **361** (1993) 77– 83.
[91] Fatouros, N., Simonin, J.P., Chevalet, J., Reeves, R.M., *J. Electroanal. Chem.* **213** (1986) 1–16.
[92] Krulic, D., Fatouros, N., Chevalet, J., *J. Electroanal. Chem.* **287** (1990) 215–227.
[93] Chevalet, J., Champagne, G.Y., Gastonguay, L., Lacasse, R., Ladouceur, M., Fatouros, N., Krulic, D., *J. Chim. Phys.* **93** (1996) 804–817.
[94] Menon, V.P., Martin, C.R., *Anal. Chem.* **67** (1995) 1920–1928.
[95] Ugo, P., Moretto, L.M., Bellomi, S., Menon, V.P., Martin, C.R., *Anal. Chem.* **68** (1996) 4160–4165.
[96] Amatore, C., Saveant, J.M., Tessier, D., *J. Electroanal. Chem.* **147** (1983) 39–51.
[97] Sabatani, E., Rubinstein, J., *J. Phys. Chem.* **91** (1987) 6663–6669.
[98] Baker, W.S., Crooks, R.M., *J. Phys. Chem. B* **102** (1998) 10041–10046.
[99] Possin, G.E., *Rev. Sci. Instrum.* **41** (1970) 772–774.
[100] Williams, W.D., Giordano, N., *Rev. Sci. Instrum.* **55** (1984) 410–412.
[101] Schoenenberger, C., van der Zande, B.M.I., Fokkink, L.G.J., Henny, M., Schmid, C., Kruger, M., Bächtold, A., Huber, R., Birk, H., Staufer, U., *J. Phys. Chem. B* **101** (1997) 5497–5505.
[102] Routkevitch, D., Bigioni, T., Moskovits, M., Xu, J.-M., *J. Phys. Chem.* **100** (1996) 14037–14047.
[103] Zheng, M., Li, G., Zhang, X., Huang, S., Lei, Y., Zhang, L., *Chem. Mater.* **13** (2001) 3859–3861.
[104] Chailapakul, O., Sun, L., Xu, C.J., Crooks, R.M., *J. Am. Chem. Soc.* **115** (1993) 12459–12467.
[105] Che, G.L., Cabrera, C.R., *J. Electroanal. Chem.* **417** (1996) 155–161.
[106] Jeonng, E., Galow, T.H., Schotter, J., Bal, M., Ursache, A., Tuominen, M.T., Stafford, C.M., Russel, T.P., Rotello, V.M., *Langmuir* **17** (2001) 6396–6398.
[107] Cheng, J.F., Whiteley, L.D., Martin, C.R., *Anal. Chem.* **61** (1989) 762–766.
[108] Hulteen, J.C., Menon, V.P., Martin, C.R., *J. Chem. Soc. Faraday Trans.* **92** (1996) 4029–4032.
[109] Daniele, S., Ugo, P., Bragato, C., Mazzocchin, G.A., *J. Electroanal. Chem.* **418** (1996) 29–34.
[110] Brunetti, B., Ugo, P., Moretto, L.M., Martin, C.R., *J. Electroanal. Chem.* **491** (2000) 166–174.

List of Abbreviations

Abbreviation	Explanation
ASV	anodic stripping voltammetry
CSV	cathodic stripping voltammetry
DL	detection limit
IEV	ion-exchange voltammetry
LB	Langmuir-Blodgett
MSW-DD	double differential multiple square-wave voltammetry
MSWV	multiple square wave voltammetry
NEAs	nanoelectrode arrays
NEEs	nanoelectrode ensembles
PVP	poly(4-vinylpyridine)
QPVP	quaternized poly(4-vinylpyridine)

PART 2

Sensor Applications

2.1 Modular Microoptical Systems for Sensors and Telecommunication

U. Wallrabe and J. Mohr, Forschungszentrum Karlsruhe, Karlsruhe, Germany

Abstract

This paper focuses on fabrication and assembly processes of modular and hybrid microoptical systems in use in sensor and telecommunication fields. These microoptical systems are made by a huge variety of processes, which are commonly known as microsystem technology. Nevertheless, quite a lot of them are based on the LIGA process, since LIGA is an "assembly friendly" technology.

Modular microoptical systems typically comprise several functions incorporating mechanical or electromechanical components beside the optical elements. The assembly of the diverse submounts and components is performed actively as well as passively. In many cases, all of these functions can be combined on one single chip. In several cases, however, a pair of chips, which is assembled in a flip-chip like way, is used.

The framework of the paper leads to a definition of the terminus "modular microoptical system" and argues for the advantages of such a modular design approach. It closes with some remarks on automatic assembly strategies.

Keywords: microoptics; modular; hybrid; flip-chip; assembly

Contents

2.1.1 Introduction 144
2.1.2 Definition of a Modular Microoptical System 145
2.1.3 Optomechanical Systems 146
2.1.3.1 Fiber Coupler in Silicon with Rhombus-like Grooves 147
2.1.3.2 Easy-assembly Fiber Coupler from Polymer 148

2.1.3.3 Expanded Beam Fiber Coupler 149
2.1.4 Opto-electromechanical Systems 1: Optical Benches Combined with Electrooptical Components 151
2.1.4.1 Heterodyne Receiver 151
2.1.4.2 Spectrometer 153
2.1.4.3 Distance Sensor 154
2.1.4.4 Blood Flow Sensor 155
2.1.5 Opto-electromechanical Systems 2: Optical Benches Combined with Electrooptical Components and Actuators 157
2.1.5.1 Segmented Deformable Mirror 158
2.1.5.2 Optical Cross Connect with Linear Motion Actuated by Electrostatics . 159
2.1.5.3 Optical Cross Connect with Vertical Motion Generated by Piezo Actuators 160
2.1.5.4 Optical Cross Connect with Rotating Mirrors Actuated by Electrostatics 162
2.1.5.5 Oscillating Modulator for Infrared Light Actuated by Electromagnetics 164
2.1.5.6 Laser Scanner for Barcode Reading Actuated by Electromagnetics . . . 166
2.1.6 A Final Word About Assembly Technology 167
2.1.7 Conclusions 172
2.1.8 References 173

2.1.1 Introduction

The tremendous increase in demand for rapid information and communication has pushed light to the number one data transfer medium. The great advantages of light are not only its unique velocity but also its insensitivity to disturbance because the photons do not interact either with external fields or with themselves. This makes light a very attractive medium to transfer information, especially across long distances but also on short stretches. The percentage of information that is already transferred by light through optical fiber networks differs from country to country. In Germany, for example, it has already reached 90% since investigations started quite early in this promising technology [1]. Recent interest in optical telecommunications lies not only in global networks, but also in private households [2] and in cars [3]. Hence optical components that enter the consumer market are subject to increased price pressure.

Similar arguments are valid for optical sensors. Again, the insensitivity to disturbance is the main reason for the increasing application of optical sensors in cars and industrial production for control and monitoring of automatic processes. Also in medical technology, optical sensors are of great interest because they are mostly non-invasive and can be used in the proximity of strong fields, eg, close to the magnetic field of a nuclear resonance tomograph [4].

Striving for cost reduction through higher integration and on-going miniaturization, the field of micro optics becomes an enabling technology for optical tele-

communication and sensors. There are two basic technology approaches: a fully integrated fabrication sequence or a modular or hybrid set-up. The choice of which one is the most suitable is related to the number of pieces to be fabricated, the complexity of available processes and the respective yield, and the freedom in design and construction. Each approach has its specific advantages. In case of lower numbers and for the combination of several functional components, the modular design and fabrication is often the better choice. The processes are easier and safer, and one can work with pre-tested components, which helps to enhance the yield.

This chapter concentrates on micro optical systems for telecommunication and sensors that follow the modular approach, in addition based on micro electro mechanical system (MEMS) or micro opto electro mechanical system (MOEMS) techniques. Examples for diverse purposes will be presented.

2.1.2 Definition of a Modular Microoptical System

The definition of the expression 'modular microoptical system' as it will be understood in this paper is best done in the reverse direction. An 'optical system' comprises a variety of functions in one unit: optical function and normally electrical and/or mechanical function. Examples of pure optical function are simply light transfer, refraction and diffraction, imaging, filtering, light splitting or superposition. Electrical or optoelectrical function can be light generation, detection or analysis, electronic circuitry, or current/voltage supply to other functional components. Mechanical or electromechanical function is provided by alignment structures or positioning aids, fixing and clamping structures, actuators, heaters and coolers.

Such an optical system is called 'micro' when at least one component of it is fabricated by means of micro-technology. The overall size of the system or even of the diverse components is not necessarily in the range of micrometers, but maybe some structural detail calls for micro-technology. A huge variety of such microoptical systems have been presented in recent years. These can simply be fiber couplers [5] or beam splitters [6], or more complex devices as switchable mirrors [7, 8], fiber switches [9, 10], Fabry-Perrot interferometers [11], or modulators [12].

A microoptical system is defined as 'modular' or sometimes just 'hybrid' if the diverse functions are not fabricated in one common fully integrated fabrication sequence and/or if they are not integrated on the same substrate. An optical bench, eg, fabricated with an adapted process on one substrate, and an actuator, which is fabricated with another specific process on a second substrate, form two modules, which are finally put together. In addition, the components required for each module can be fabricated separately and mounted in a hybrid way. It is clear that the definition of 'modular' and 'hybrid' cannot be used rigidly. In gen-

eral, the individual scientist decides whether he or she considers a system just developed already to be a modular or just a hybrid set-up. Also in this chapter, we did not succeed in separating them. Hence modular and hybrid systems will necessarily be mixed up or even lead to contradictory points of view. We will stress this aspect in Section 2.1.5, which describes the most complex modular systems in this chapter.

Anyhow, no matter what name it is given, a modular or hybrid microoptical system is always related to assembly technology comprising alignment and fixation [13]. This can be performed either actively or passively. Active alignment means that the output signal of the system is monitored during assembly in order to obtain an optimum result. Passive assembly, in contrast, makes use of integrated alignment aids for the diverse components without the need for monitoring. The question of which method is the most suitable needs to be answered for each system separately. It is dependent on the function of the system and on the available fabrication and assembly procedures.

The examples of modular microoptical systems that are given in the following sections are split into three categories. The first covers pure opto-*mechanical* systems without any need for electrical units. The second covers opto-*electro*-mechanical systems in which the electrical units are responsible for light generation and detection. The third also describes opto-electro-mechanical systems, but including electrically driven actuators. In order not to overload this chapter, only the modular concepts and operating principles of the systems will be explained. For detailed results we refer to the original publications.

2.1.3 Optomechanical Systems

Pure optomechanical modular systems can hardly be found. Again, the boundary between a monolithic, integrated system and a hybrid, modular system is floating. This becomes clear from consideration of an optical bench with non-integrated but assembled microoptical components. The base plate of an optical bench mostly is a monolithic device with alignment aids and perhaps some integrated devices such as a prism or a mirror. When additional components such as fibers, lenses, beam splitters, etc., are assembled on the bench it becomes a system according to the definition above. A very simple example is a fiber coupler as used for telecommunication purposes with the optical function of guiding light from one fiber to the next and the mechanical function of alignment and fixation. Three different approaches to fiber couplers are presented in the following discussion.

2.1.3.1 Fiber Coupler in Silicon with Rhombus-like Grooves

Fibers are commonly aligned by grooves on a chip running parallel to the fiber axis or they are pushed through holes in a chip mounted perpendicular to the fiber axis. Most popular are V-grooves etched in single crystal silicon [14]. The fibers are fixed by glue and covered with a glass plate. A newer approach is to use rhombus-shaped channels [15] as indicated in Figure 2.1.1. The grooves allow for lateral and vertical positioning of the fibers. The authors claim two advantages over simple V-grooves. First, the volume of the rhombus is smaller, thus reducing the amount of glue which is necessary. This again reduces shear stress in the whole assembly of the Si chip, the cover glass plate, the fibers, and the glue owing to the different thermal expansion when temperature changes occur. Second, fiber assembly is very easy because funnels are generated owing to the etching process at the edges of the chip. The fibers are inserted at the wide end

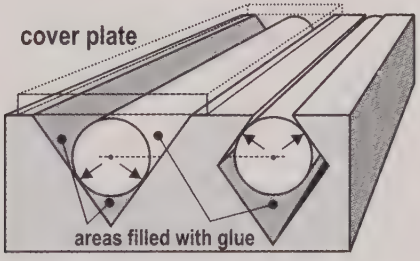


Figure 2.1.1. Comparison of fiber grooves with triangular and rhombus-like shapes [15] (with permission of Springer-Verlag).

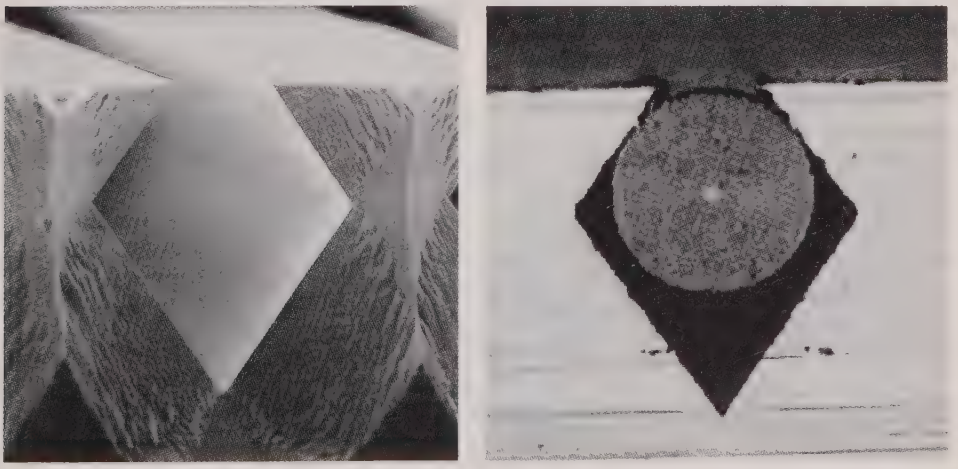


Figure 2.1.2. Rhombus-like channel with funnel at the end etched in silicon and fiber assembled in channel [15] (with permission of Springer-Verlag).

of the funnels, and as they are pushed forward they center by themselves (Figure 2.1.2). The end face is finally prepared with a wafer saw.

2.1.3.2 Easy-assembly Fiber Coupler from Polymer

A coupler with longitudinal grooves for 16 multimode fibers is described in [16], and is illustrated schematically in Figure 2.1.3. It consists of two plastic pieces made from poly(methyl methacrylate) (PMMA): one for the alignment of fibers and guide pins with five rows of highly precise alignment structures, and the other for their fixation and protection. Elastic ripples in the side walls of the alignment structures, as can be seen in a scanning electron microscope (SEM) photograph in Figure 2.1.4, facilitate fiber and pin insertion. They also make the connector insensitive to variations in fiber diameter. The gap between the alignment structures decreases successively from the last row to the first row facing the opposite connector. This permits very easy assembly and passive alignment of the fibers without the need for micro-positioning. Since the gap at the front face is 2 μm smaller than the fiber diameter, the fibers are clamped gently by the alignment structures, thus allowing easy handling during the on-going assembly. To bond both parts, UV-curing adhesive is filled into the device through a hole in the upper substrate. The glue spreads into the coupler just by capillary forces. Finally, the front face is polished. The pins of both connectors to be coupled are inserted into a highly precise coupling adapter as indicated in Figure 2.1.5. If the pins damage the respective holes in the adaptor owing to repeated reconnection, the adaptor can be replaced easily whereas the fibers and connectors can be maintained. The coupler is fabricated by means of plastic micro-injection molding made from PMMA. The respective molding tool was fabricated by LIGA technology and micro-precision engineering as described in detail in [16].

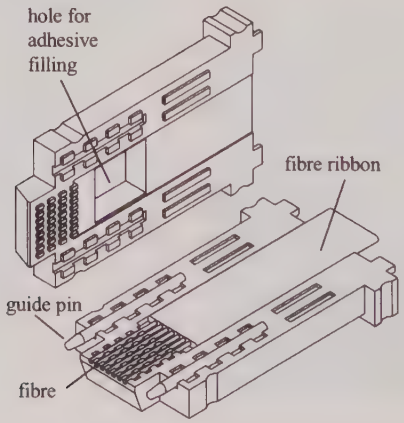


Figure 2.1.3. Schematic of ferrule.

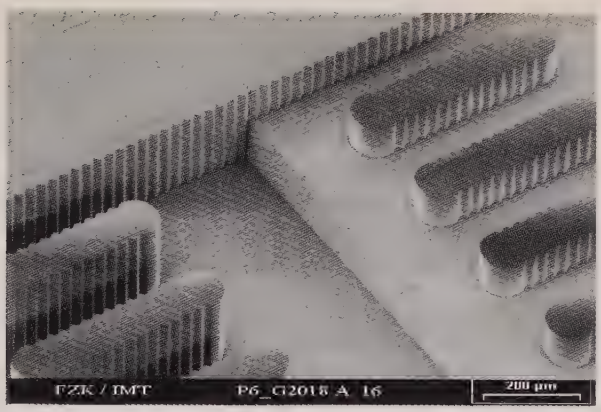


Figure 2.1.4. SEM of alignment structures.

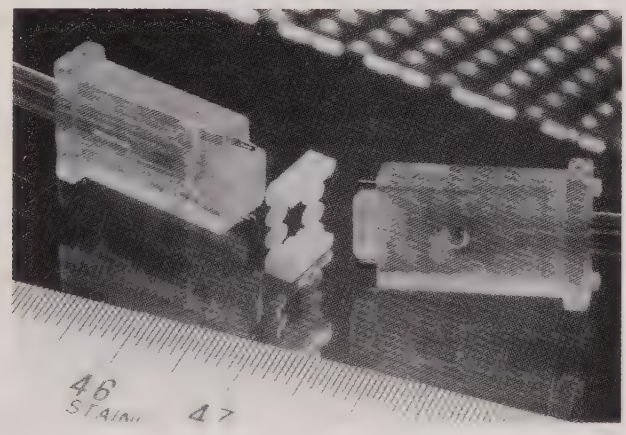


Figure 2.1.5. Pair of ferrules with coupling adapter.

2.1.3.3 Expanded Beam Fiber Coupler

A comparable connector concept is reported in [17]. The authors also describe an injection molded coupler for which the tool was made in a similar way; however, the precision fiber grooves are obtained by V-grooves etched in silicon. They extended this concept to an expanded beam coupler [18] with the advantage of decreased sensitivity to fiber misalignment and to impurities. In that case the fibers are not coupled face to face but lens arrays are placed between them in order to expand the light beam for coupling and refocus it again into the output fibers as sketched in Figure 2.1.6. For this purpose, lens alignment structures are positioned in front of the V-grooves. Figure 2.1.7 shows the front face of

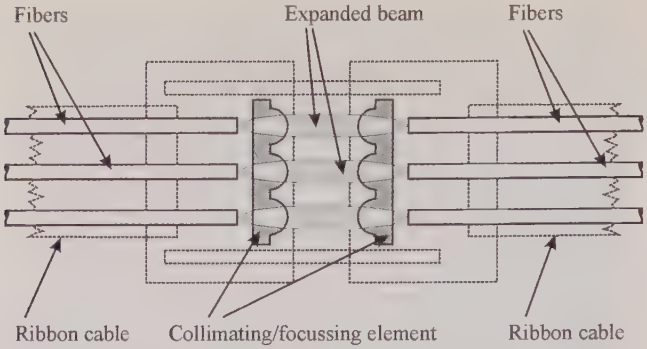


Figure 2.1.6. Principle of an expanded beam fiber coupler [18] (with permission of SPIE).

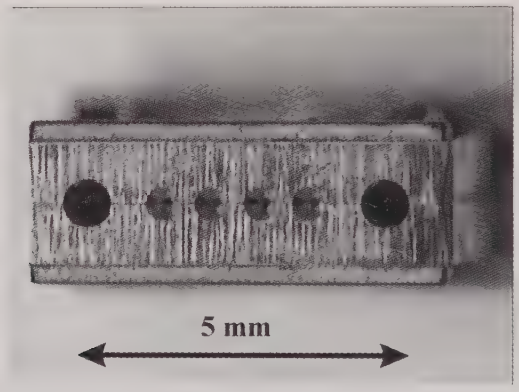


Figure 2.1.7. Front view of a molded expanded beam coupler with alignment structures for a lens array [18] (with permission of SPIE).

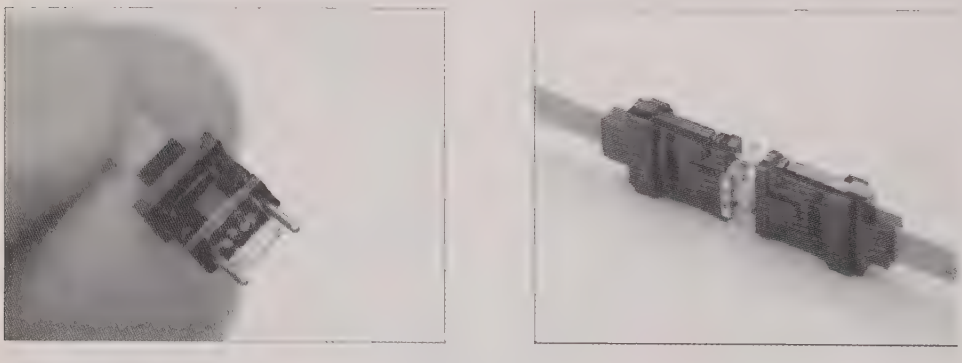


Figure 2.1.8. Assembled coupler with lens array, guide pins, and fiber ribbon [18] (with permission of SPIE).

such a micro-molded coupler. The four smaller circular holes are the place holders for the lenses and the two bigger holes for the guide pins. The coupler is equipped with fiber ribbon, a lens array of 500 µm diameter lenses and guide pins as shown in Figure 2.1.8.

2.1.4 Opto-electromechanical Systems 1: Optical Benches Combined with Electrooptical Components

The combination of electrooptical functions given by light sources and detectors with optical benches as described above follows more the idea of a microoptical *system.* The totally different purposes of opto mechanics and opto electronics call for a modular approach since the integration of all functions on a monolithic substrate makes processes very complex, difficult, and expensive. Thus, examples of this category of modular microoptical systems are found in a larger variety. Nevertheless, reviewing the literature one finds that this field is dominated by LIGA-made systems. This may be related to the fact that, on the one hand, thick microstructures are helpful for alignment, and on the other hand, groups working on silicon processes and optical lithography tend to look for a higher integration level. Therefore, we present three systems made by the LIGA technique but close this section with a sensor system made by silicon micromachining and optical lithography.

2.1.4.1 Heterodyne Receiver

Figure 2.1.9 shows the set-up of a heterodyne receiver, ie, a wavelength filter for telecommunication. In this case, two incoming light beams need to be split and superposed again [19]. The signal light and the light from a local laser are coupled into the system by means of monomode fibers. The light is collimated by ball lenses and then split into its two polarization states. Reaching the next optical surface, the beam of each polarization state is again split by 50% and is simultaneously superposed with the respective beam from the opposite light source. Each of the final four superposed beams is detected by a photodiode. The system consists of a ceramic chip on which alignment structures from the polymer are patterned using LIGA technology. The fibers, the ball lenses, the prisms for the beam splitters, and the diodes are separately manufactured components, which are assembled in a fully passive way on the chip. They are just pushed towards the alignment structures and fixed with some UV-curing glue. The accuracy is of the order of 1 µm. Since the altitude of the optical axis is defined by the diameter of the ball lens (here

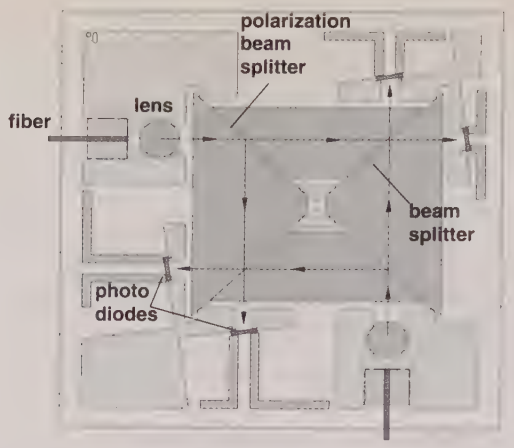


Figure 2.1.9. Set-up of heterodyne receiver.

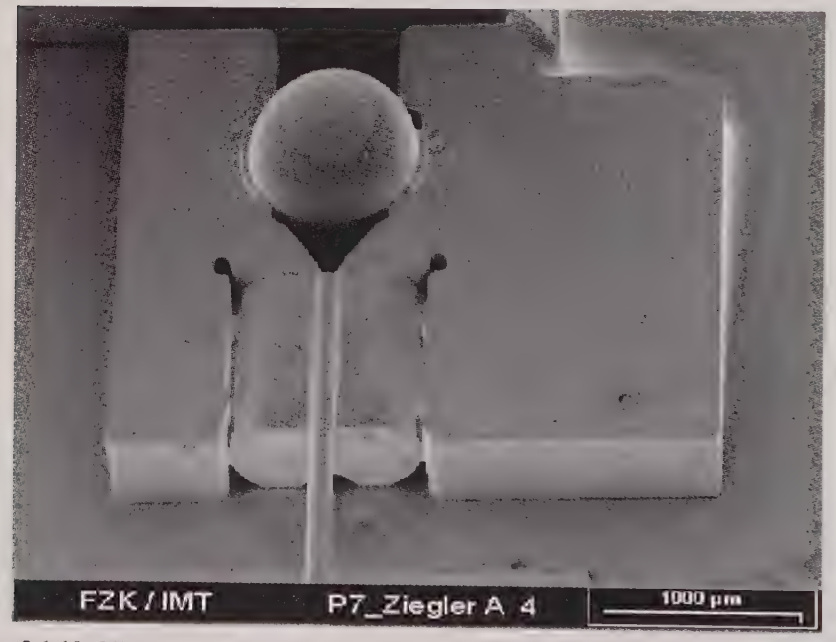


Figure 2.1.10. Fiber mount levelling the fiber on the optical axis defined by the ball lens.

900 µm), a fiber with a diameter of only 125 µm needs to be levered to the same height. This is obtained by a fiber mount as becomes clear from Figure 2.1.10. The electrical connection of the diodes is done with gold tracks, which were pre-patterned on the substrate by optical lithography and wet etching.

2.1.4.2 Spectrometer

The LIGA-made microstructures of the heterodyne receiver have a pure mechanical, ie, alignment function. However, they may also have a very distinct optical function as is the case with micro-spectrometers [20], which are illustrated in Figure 2.1.11. Here, white light is coupled into the system by a multimode fiber, which is put in a fiber alignment groove. The light spreads into the free space, forming a hollow wave guide by Fresnel reflections. The light is diffracted at a

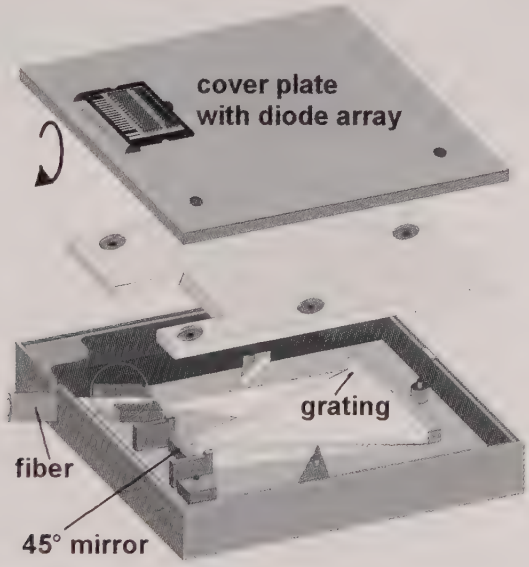


Figure 2.1.11. Schematic of spectrometer.

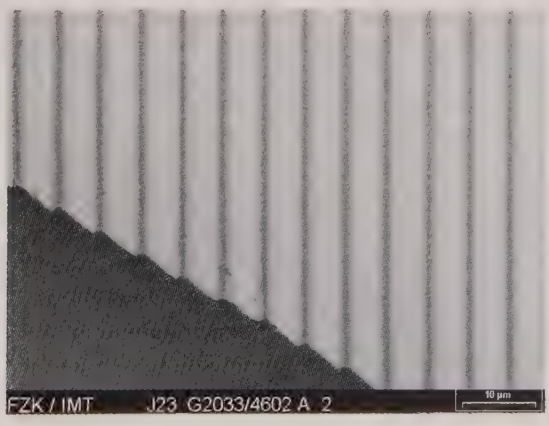


Figure 2.1.12. SEM of grating of molded IR spectrometer.

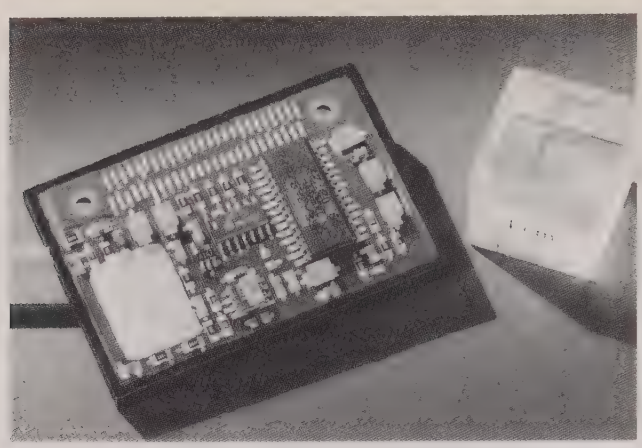


Figure 2.1.13. Full spectrometer system.

reflection grating (Figure 2.1.12), and is coupled out of the system at a 45° mirror next to the fiber. The fiber groove, the grating, and the 45° mirror are fabricated in one step by hot embossing techniques [21] using a molding tool made by LIGA technology. The polymer typically used for hot embossing is PMMA. Hence a gold layer needs to be sputtered on the PMMA structures to make them reflective. A cover plate seals the spectrometer and serves also as the upper surface of the hollow wave guide. An electrooptical chip with a photodiode array and analog electronics is finally put head-over on top of the spectrometer. The full system with the size of a match box is displayed in Figure 2.1.13.

2.1.4.3 Distance Sensor

A similar idea of a hybrid assembly was used for a micro-distance sensor working on the basis of the triangulation principle [22], ie, the position of the spot on the sensor's detector is dependent on the distance of the object to be measured. As with the spectrometer, the distance sensor is divided into two functional units: a passive optical chip again fabricated by LIGA technology including bent mirrors, 45° mirrors, and alignment structures for cylindrical lenses (Figure 2.1.14a), and an electrooptical chip (Figure 2.1.14b) with a laser, a photodiode and a position-sensitive detector (PSD). The optical components on the LIGA chip are made from PMMA covered with an evaporated gold layer. Both chips are again mounted head-over to form an opto-electro-mechanical sensor system. For easy assembly, two micro-spheres of glass are used as position aligners of the LIGA and the Si chip (Figure 2.1.14c). First, the micro-spheres are placed into pyramidal grooves on the Si base, thus centering themselves. Then, the optical chip with two cylindrical holes for the spheres is passively aligned on the Si chip. The two chips before the head-over assembly are shown in Figure 2.1.15. The Si chip is already mounted on a carrier in a TO8 housing.

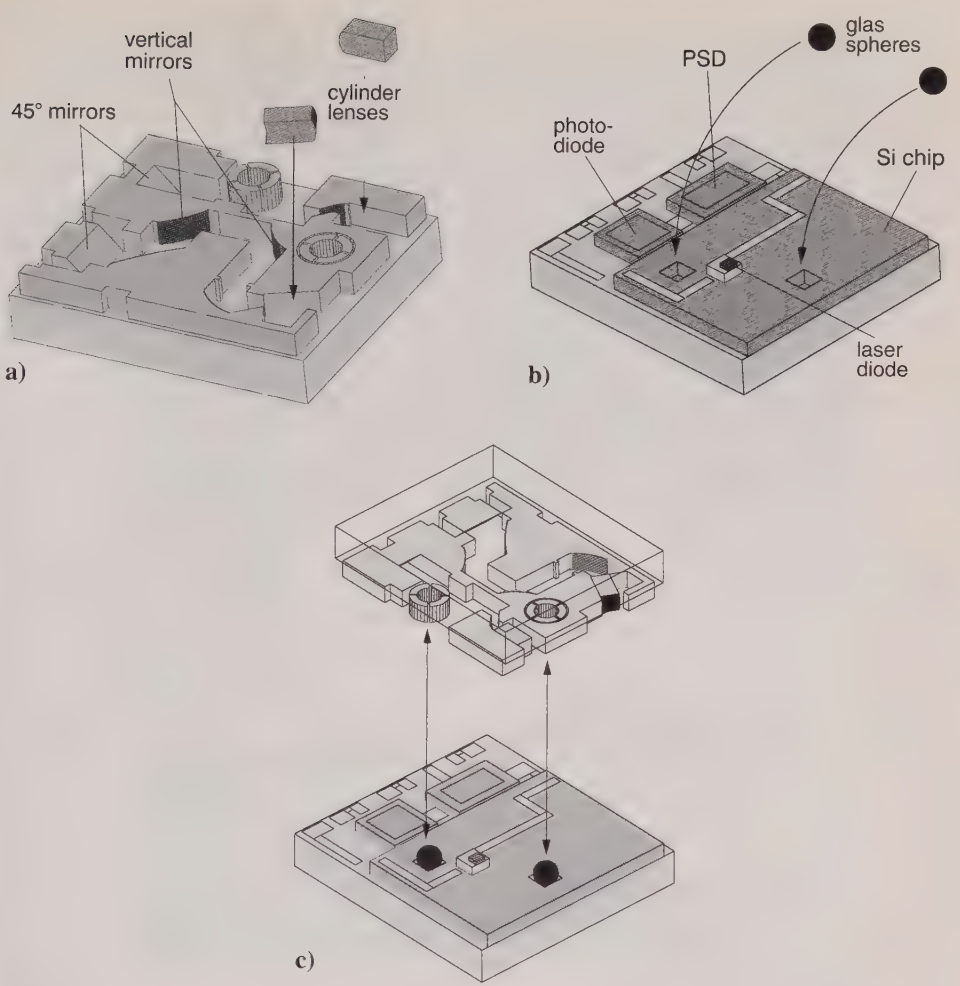


Figure 2.1.14. (**a**) Passive optical chip; (**b**) electrooptical Si chip; (**c**) head-over assembly of opto-electro-mechanical distance sensor system [22] (with permission of Springer-Verlag).

2.1.4.4 Blood Flow Sensor

Not only LIGA structures provide the advantage of optical function in parallel with alignment accuracy. Another good example is a non-invasive blood flow sensor that makes use of the Doppler method [23]. Once again, the device is an assembly of two chips [24], as can be seen in Figure 2.1.16: one Si chip with V-grooves serves as a protection cap for the lower chip, which is an Si substrate with a polyimide waveguide, a laser diode (LD), a photodiode (PD) and electrodes on top of it. The Si substrate was patterned with terraces, Au/Pt/Ti elec-

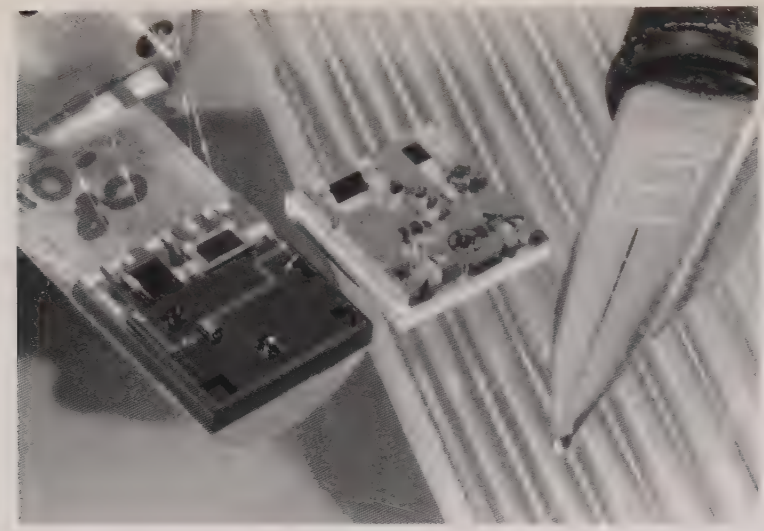


Figure 2.1.15. Photograph of two chips before final assembly [22] (with permission of Springer-Verlag).

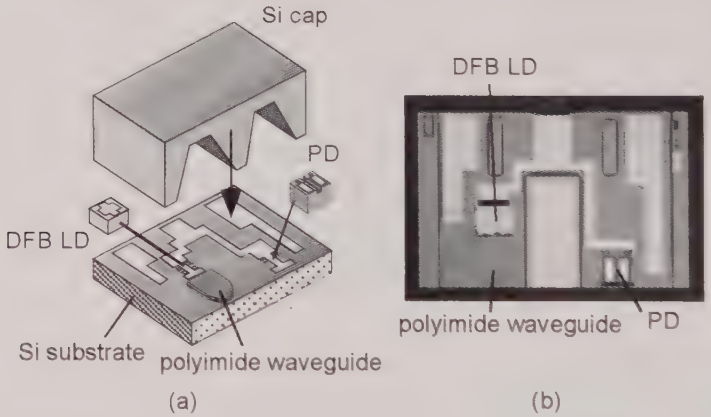


Figure 2.1.16. Schematic set-up of blood flow sensor and photograph of the Si substrate [24] (with permission of IEEE).

trodes, AuSn solder layers, and a three-layer polyimide waveguide. The convex-shaped edge of the waveguide collimates the light from the edge-emitting laser in the horizontal direction, thus allowing operation without the need for an additional lens. Figure 2.1.17 shows a close-up view of the front face of the chip package and the fully assembled sensor system, which is only 20 mm long.

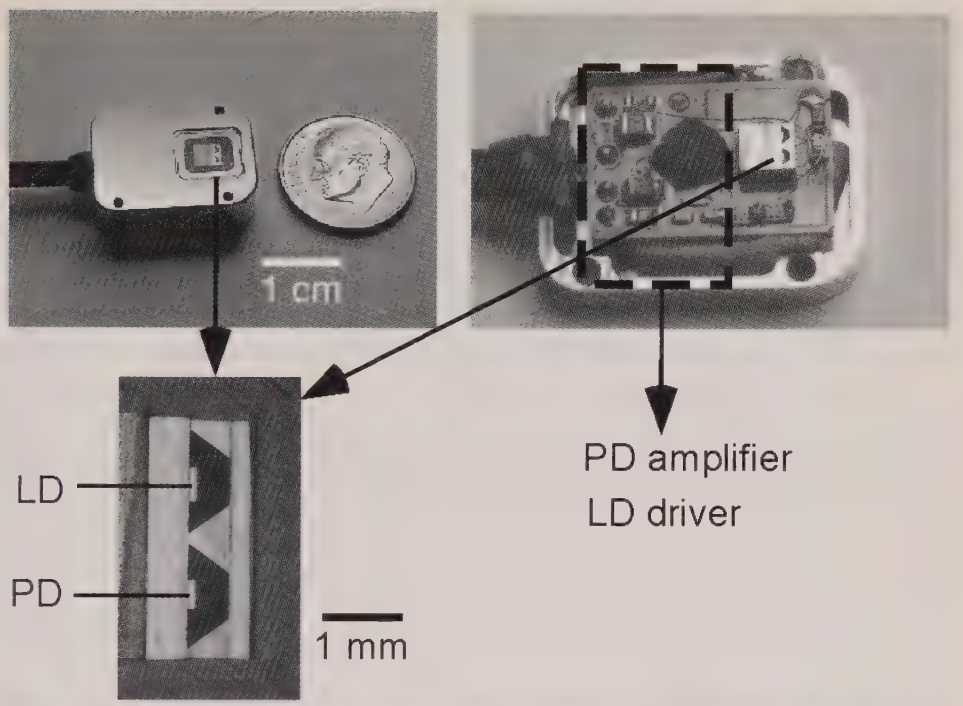


Figure 2.1.17. Close-up of front face of chip package and fully assembled sensor [24] (with permission of IEEE).

2.1.5 Opto-electromechanical Systems 2: Optical Benches Combined with Electrooptical Components and Actuators

In contrast to the preceding section, the idea of a system is now extended by the inclusion of actuators. Again, a modular set-up has the advantage of much easier processes and gives rise to more degrees of freedom for linear movement, rotation, and tilting. Also, the kind of actuation is easier to choose among electrostatics, electromagnetics, piezoelectric or thermal effects, or even hydraulic pressure or capillary forces. The choice of actuation principle is dependent on the required displacement or angle as a function of the force, on boundary conditions such as the available driving current or voltage, the allowed temperature rise, or the environmental conditions in the working area.

This category of opto-electro-mechanical systems is dominated by applications from the telecom field. Although a huge number of publications are concerned with fully integrated systems, there are also some dealing with a modular set-up. No matter whether scanners or switches for fiberoptical networks are discussed, the main focus is put on steerable mirrors. The six systems presented in this sec-

tion exemplify the variety of kinds of motion and actuation principles. In addition, they represent diverse levels of the modular approach. As already mentioned in the Introduction, we will use the modular microoptical systems of this section to discuss the definition 'modular' once more, now with practical examples.

2.1.5.1 Segmented Deformable Mirror

As in the previous section, a flip-chip-like assembly concept was used for a segmented deformable micromirror array for free-space optical communication as is explained in Figure 2.1.18 [25]. The upper chip contains on its top surface segmented mirrors, which are attached to actuator segments, each of which is suspended from four cantilevers in order to pull them down. The lower chip contains square electrodes to apply electrostatic forces to the actuators and supporting side walls in order to bond the two chips mechanically. All functional elements, ie, the mirrors, actuators, supporting side walls and electrodes, are patterned in polysilicon. For thermal compression bonding, gold bond pads are deposited on top of the side walls and onto the frame of the mirror segments. The spacing between the actuators and electrodes is about 5 μm, which results in displacements of the mirrors of 1 μm at a voltage of 200 V. An SEM photograph of one pixel of the assembled device is shown top-down in Figure 2.1.19. Part of the mirror has been broken off to show the underlying actuator and electrodes.

This concept easily allows one to create large mirror arrays such as 1000×1000 pixels and to put the electrodes directly on integrated circuitry. Thinking on such a big scale, a single mirror is just a component of the whole

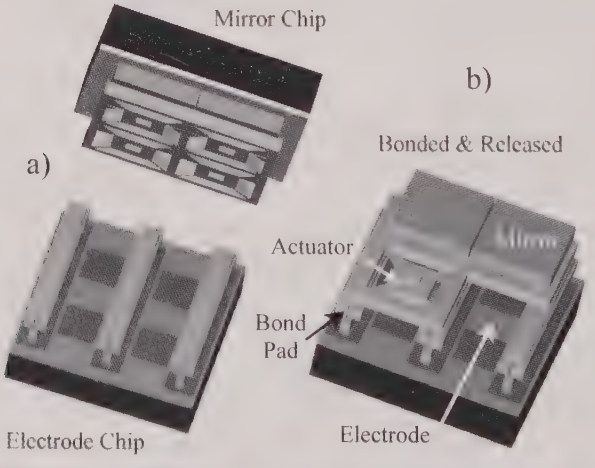


Figure 2.1.18. Schematic 3D drawing of a deformable micro-mirror array before and after bonding. The vertical scale is expanded 10-fold [25] (with permission of IEEE).

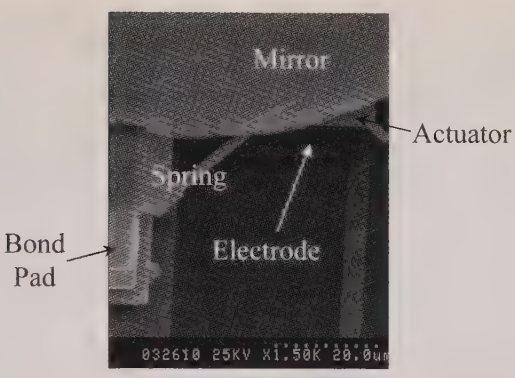


Figure 2.1.19. SEM of a single pixel [25] (with permission of IEEE).

system. However, on a small scale the single mirror can also be considered as a system made of two modules: an opto mechanical (upper chip) and an electrical (lower chip) module.

2.1.5.2 Optical Cross Connect with Linear Motion Actuated by Electrostatics

Again, flip-chip assembly was chosen for an optical fiber cross connect with MEMS switches [26] and integrated waveguides [27]. Figure 2.1.20 shows the MEMS fiber switch made by deep reactive ion etching (DRIE). It consists of two linear electrostatic comb-like micro actuators, which generate a linear movement of the long bar in the center. The mirror (top left corner in Figure 2.1.20), which resembles a needle in this design, is part of the bar and can move into the free space between the grooves for the waveguides. Sixteen of these switches are arranged on the substrate to form a 4×4 switch array as can be seen on the left side of Figure 2.1.21. A second chip provides the integrated waveguides made from SiO_xN_y on SiO_2 also patterned by DRIE. At the position of the mirrors, the waveguides become slimmer in order to bring them as close together as possible and thus reduce optical losses (Figure 2.1.21, right). In addition, the faces of the waveguides can be given an optical shape. The waveguide chip is put head-over on the mechanical chip and fibers are attached to the edges of the package (Figure 2.1.21, center).

In this case, the modular aspect is obvious and easy to find. The two chips have absolutely distinct functions, one for optical waveguiding with integrated imaging, the other providing the electromechanical actuator. Hence the complete system is an assembly of two modules.

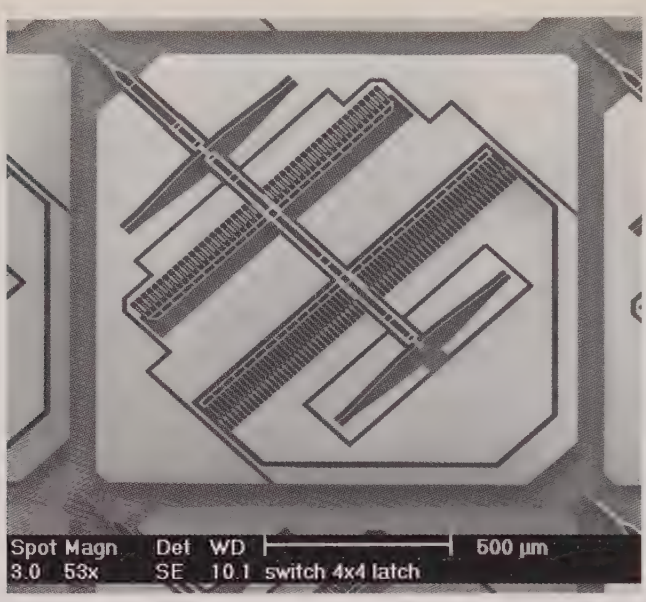


Figure 2.1.20. Flip-chip set-up of 4×4 optical cross connector [27] (with permission of IEEE).

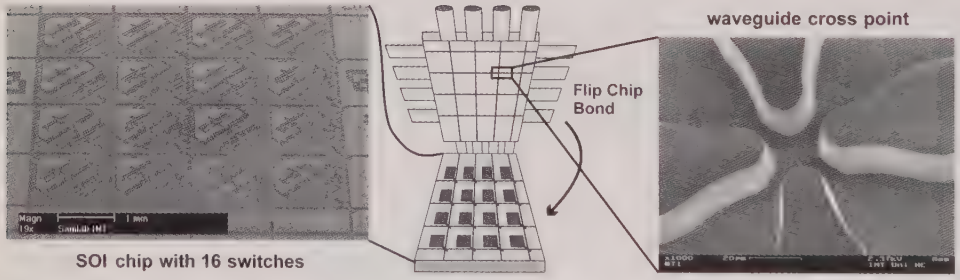


Figure 2.1.21. Close-up view of waveguides at the cross point [27] (with permission of IEEE).

2.1.5.3 Optical Cross Connect with Vertical Motion Generated by Piezo Actuators

A different approach for the realization of a 4×4 optical switch is described in [28]. It is a fully hybrid set-up of mostly commercial components. The principle becomes clear from Figure 2.1.22, which shows a 1×4 channel system. The light from the incoming fiber is collimated by miniature collimation optics. The beam passes four cantilevers with prisms on them. The cantilevers are piezoelectric bimorph actuators, which bend in the vertical direction, thus pulling the 'non-active' prisms below the optical axis. Depending on the position of the only 'ac-

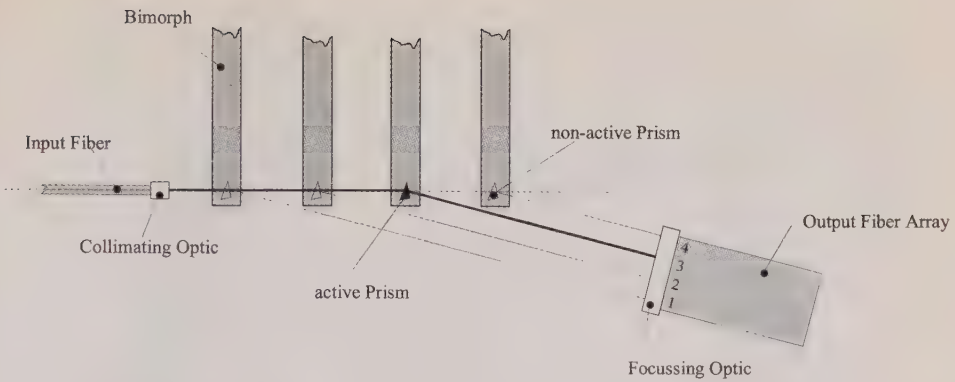


Figure 2.1.22. Schematic drawing of hybrid fiber switch based on piezo actuators [28] (with permission of SPIE).

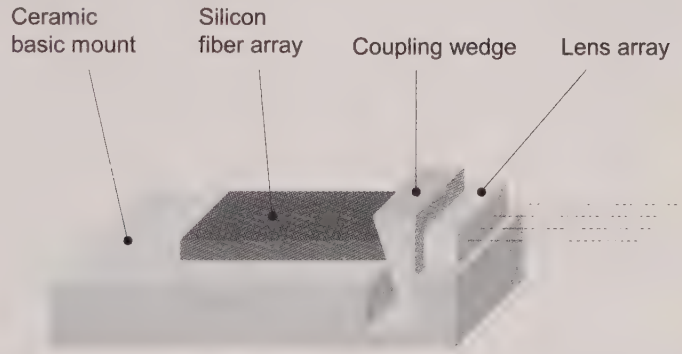


Figure 2.1.23. Schematic drawing of miniaturized focusing optics [28] (with permission of SPIE).

tive' prism, one of the four output fibers is chosen. The light is coupled into the output fibers by means of another miniaturized focusing optics. Multiplying this subsystem by a factor of four, the overall system is obtained. For assembly of the piezo actuators, a ceramic submount providing precise alignment grooves was prepared. The prisms were assembled all in parallel using a vacuum matrix gripper [29] with triangular precision engineered grooves. The miniature focusing optics are shown in Figure 2.1.23. An Si fiber array with an angle polished edge is mounted on a ceramic base. In order to minimize back reflections, a coupling wedge is used to guide the light from the fibers to a micro-lens array. The three components are actively aligned and fixed with UV-curing glue. The full system of a 4×4 switch array after electrical connection of the actuators and fiber insertion is shown in Figure 2.1.24.

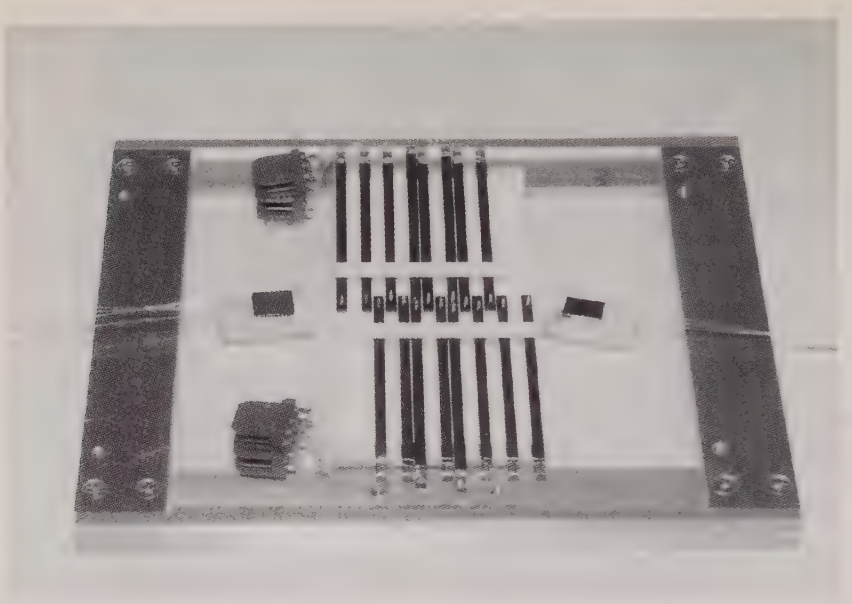


Figure 2.1.24. Fully assembled 4×4 optical switch matrix [28] (with permission of SPIE).

In this modular microoptical system, the hybrid and modular assembly aspects are mixed. The hybrid array of cantilevers with prisms represent the actuator module and the two imaging optics represent the optical modules.

2.1.5.4 Optical Cross Connect with Rotating Mirrors Actuated by Electrostatics

A third assembly strategy for a fiber switch array is to combine a LIGA-fabricated micro optical bench as described above with rotating mirrors, which are arranged in a matrix scheme [30]. Figure 2.1.25 shows the concept for an N×N switch. The light from the fibers is collimated onto the mirror surfaces and refocused into the output fibers by spherical glass lenses. The fibers and lenses are passively aligned with the help of fiber mounts and alignment stops in the optical bench. In order to achieve uniform insertion losses for all channels, the fibers are arranged in a way that all optical pathlengths are the same. The mirrors are attached to the outer side of electrostatic micro-motors, which swivel the mirrors into the beam path. Figure 2.1.26 shows the switching mechanism. Mechanical stops define the correct end position of the mirrors, which were designed as double mirrors in order to ensure always 90° reflection. The optical benches, the static motor parts, and the mirror stops are fabricated simultaneously using LIGA technology. Because of the electromechanical function of the motors, the struc-

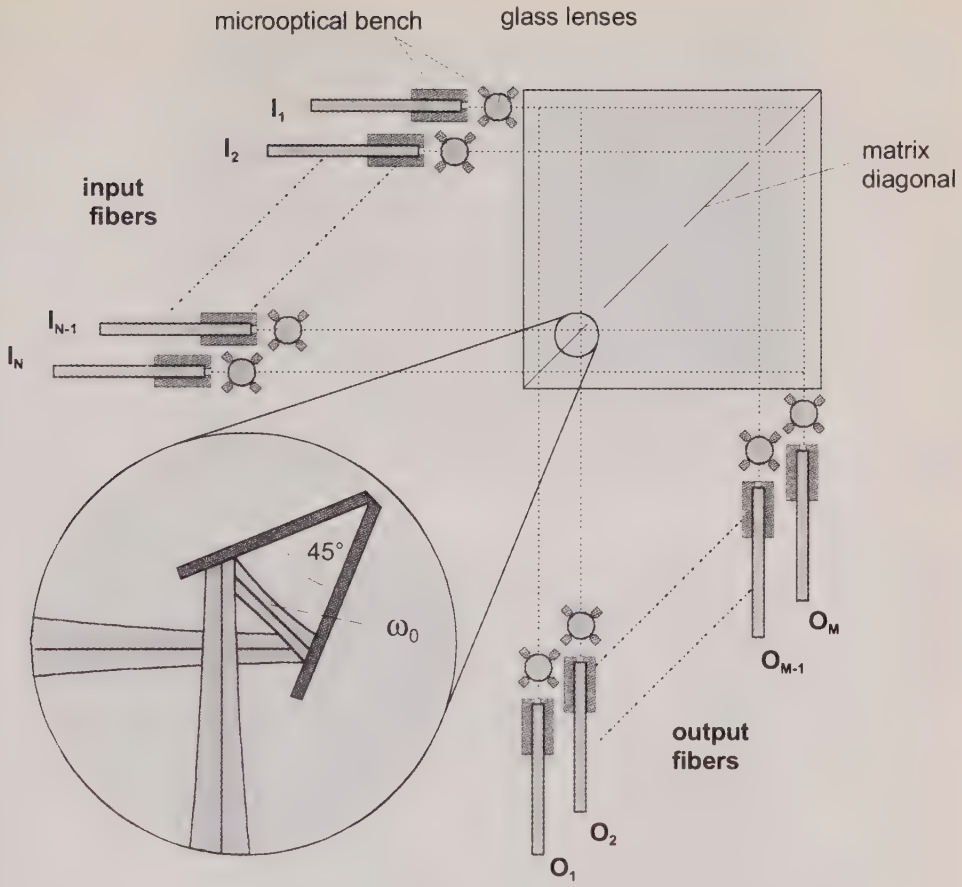


Figure 2.1.25. Schematic drawing of N×N matrix switch using rotating mirrors.

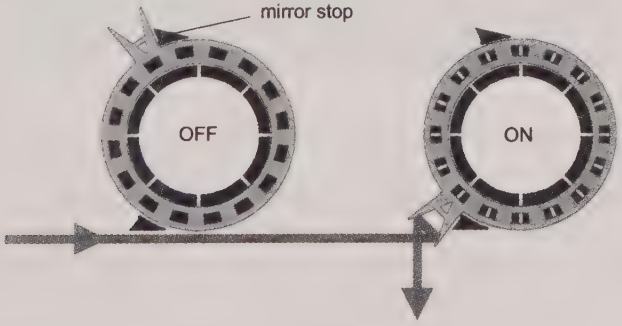


Figure 2.1.26. Switching mechanism of rotating mirrors.

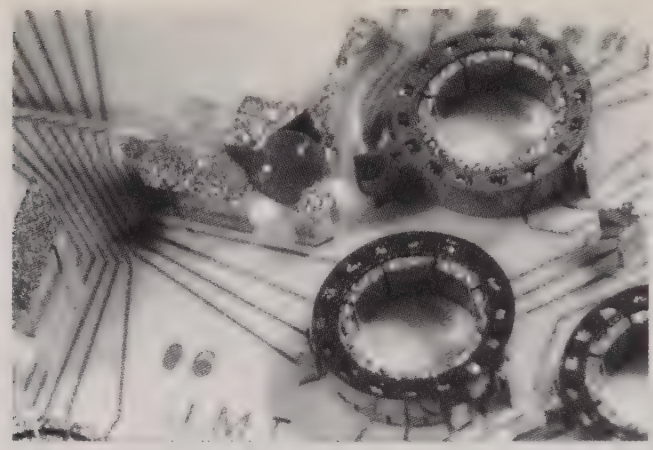


Figure 2.1.27. Close-up view of fully assembled 2×2 switch matrix.

tures are made from electroplated nickel on a sacrificial layer [31]. The rotors with the attached double mirrors are manufactured similarly on a separate substrate and are set manually above the stators. The fiber mounts from PMMA are patterned on a third substrate. Figure 2.1.27 shows a close-up view of a fully assembled 2×2 switch matrix. On the left side, the fiber on the fiber mount and a lens are pushed towards the stop structures of the optical bench. The mirror in front of the lens is in the 'on' position and the mirror below in the 'off' position.

This switch matrix can be considered as a fully hybrid set-up since all components and functions are finally placed on one substrate although they have been fabricated on three substrates in total. It is a question of preference to define the optical part with fibers and lenses as the optical module and the rotating mirror as the mechanical module.

2.1.5.5 Oscillating Modulator for Infrared Light Actuated by Electromagnetics

To complete this section about opto-electro-mechanical systems including actuators, two examples from the sensor field are given. The first is a modulator, ie, a chopper, for infrared light for suppression of noise in infrared spectrometers [32]. The working principle becomes clear from the demonstrator shown in Figure 2.1.28. Alignment structures for input and output fibers are patterned on a ceramic substrate in parallel with a movable shutter (looking like a small hammer in Figure 2.1.28) and an electromagnetic actuator, which becomes obvious from the coil in the upper part of the picture. The structures are fabricated using LIGA technology and electroplating of Permalloy, which is a soft magnetic alloy of 80% nickel and 20% iron. A current flowing through the coil generates a magnetic field, which is guided through the metallic yoke. Owing to the decrease in

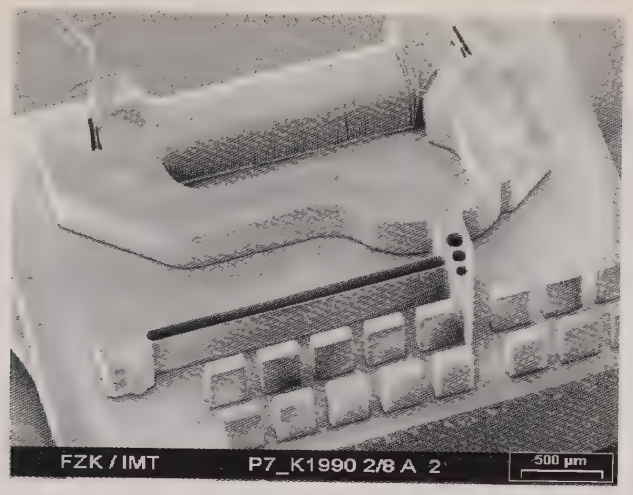


Figure 2.1.28. Demonstrator for electromagnetic microchopper.

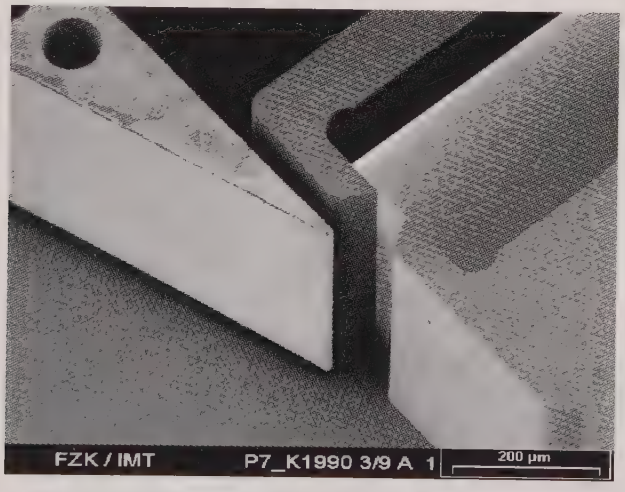


Figure 2.1.29. Close-up of shutter in front of fiber.

the resistance of the magnetic circuit, the shutter is pulled backwards between the pole shoes of the circuit, thus enhancing the metallic cross section for the magnetic field. A sinusoidal current generates a periodic field and an oscillation of the shutter. The yoke for the coil is manufactured on a separate substrate and released. The coil is wound manually so far and then assembled on the substrate using alignment and clamping structures. For integration of the chopper in a spectrometer (discussed above in Section 2.1.4), the fiber alignment structures are skipped. The chopper is put head-over in the light-input part of the spectrometer. This is facilitated by four stops on the edges of the yoke and the fixing

block of the shutter, which were patterned simultaneously for this purpose. Figure 2.1.29 shows a close-up view of the very precise position of the shutter tip directly in front of the input fiber of a spectrometer in its alignment groove.

The final system of the spectrometer including the chopper is a full modular assembly of an optical and an electromechanical module. The chopper itself is a pure hybrid device.

2.1.5.6 Laser Scanner for Barcode Reading Actuated by Electromagnetics

The second sensor example is a laser scanner for barcode reading [33] with a similar actuation principle as the chopper. Figure 2.1.30 shows a schematic drawing of the scanner. Again, a magnetic circuit with an assembled coil is fabricated from Permalloy by means of LIGA. A perpendicular mirror is attached on a freely suspended metal anchor, which is part of the magnetic circuit. The air gap in the circuit and therefore, the magnetic resistance are minimized when the an-

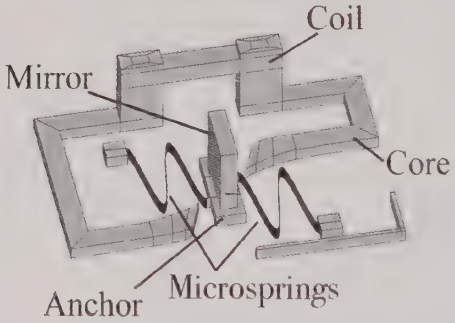


Figure 2.1.30. Schematic drawing of a laser scanner for barcode reading.

Figure 2.1.31. Stroboscopic photographs of a scanner at a deflection angle of 20°.

chor is deflected towards the ends of the pole shoes. As above, a periodic current will generate an oscillation of the mirror with twice the current's frequency. The mirror is made from a 100 µm thick Si wafer, which is coated with a reflective Au layer and cut to pieces of $1.5 \times 1.6\,mm^2$ by a wafer saw. It is placed in an alignment groove, which was patterned in the anchor and is fixed by adhesive bonding. At a current of 20 mA, the scanner reaches a scan angle of 12° in a nonresonant mode. Figure 2.1.31 demonstrates the motion with two stroboscopic photographs.

Like the chopper discussed before, the scanner is a hybrid device. In this case, however, it is a module in a 'macroscopic' optical system, which so far is not made by means of micro-technology.

2.1.6 A Final Word About Assembly Technology

No matter what categories of opto-electro-mechanical systems were discussed above, they all had one common task: assembly! Many components could be assembled passively using alignment aids such as stops or grooves, which were directly patterned into the basic structures or substrates. If not, alignment had to be done actively, which means that the performance of a device is optimized while it is operated and the components are still under positioning control. For the fabrication of demonstrators or prototypes, both assembly strategies are done more or less manually. However, for industrial fabrication, automatic assembly procedures are desired. Several research groups are concerned with that target and are developing grippers, bonders, dispensers for adhesives, and so on.

Some specific aspects for the assembly of microoptical devices are listed in [34]. The components have various but typical shapes such as spheres (ball lenses), cylinders (mirrors, GRIN lenses, fibers), plates (mirrors, lens arrays) or prisms (beam splitters). They are small, of low weight and have smooth surfaces, which enhances adhesion. They are sensitive to electrostatic fields or mechanical stress, the material typically is brittle so careful handling with the necessity of sensing small gripping forces is required. Finally, the most important goal is not to damage the optical surfaces of the components that have to be picked and placed. These handling conditions are also valid for all items concerned with storage in magazines and transportation. Some examples given in [34] are a magazine for cylinder lenses (Figure 2.1.32), a gripper for active alignment of prisms (Figure 2.1.33) and a fiber gripper and manipulator (Figure 2.1.34). The magazine allows safe storage of the cylinder lenses in a reproducible position. It provides a defined interface to the other assembly equipment with sufficient space between the lenses to pick them up. The fiber manipulator consists of two grippers, the first for stress relief and the second to pick the fiber at its uncoated tip and for fine positioning, eg, when the fiber is threaded through a hole as indicated in Figure 2.1.34.

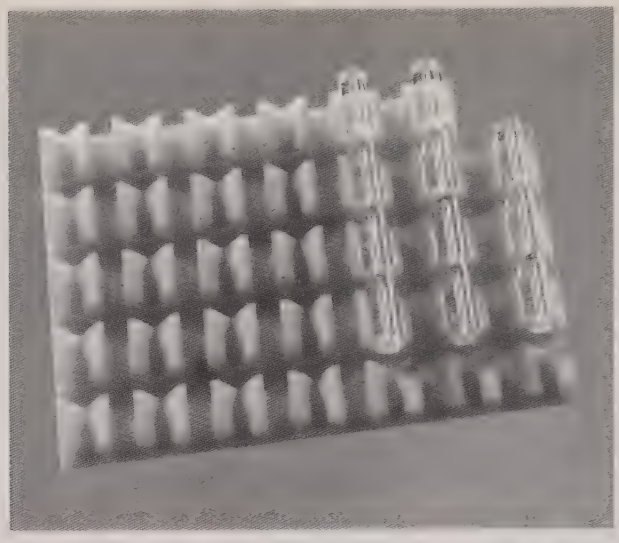


Figure 2.1.32. Magazine for the storage of cylinder lenses [34] (with permission of SPIE).

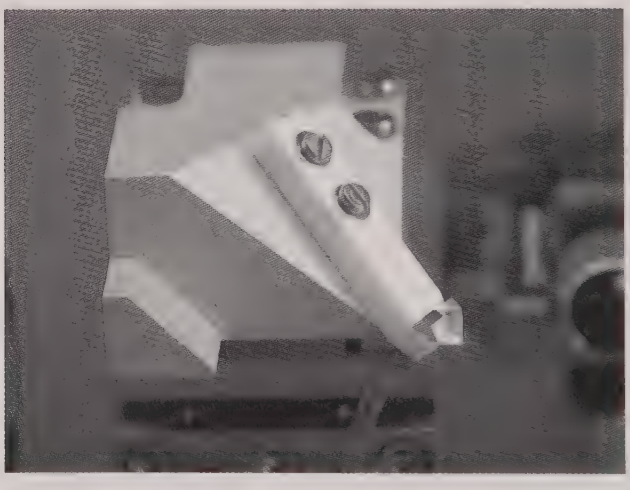


Figure 2.1.33. Gripper and alignment system for prisms [34] (with permission of SPIE).

A complete system for automatic assembly of a microoptical duplexer [35] for which ball lenses, monomode fibers, and wavelength filters have to be assembled into an optical bench and a housing is described in [36]. The authors have developed magazines and grippers for all components. All grippers work by vacuum suction except the one for the fibers. For proper material flow during assembly, a complete tray on a *xy*-stage was developed which carries all magazines as sketched in Figure 2.1.35. The tray is pushed towards three reference pins to ensure a precise

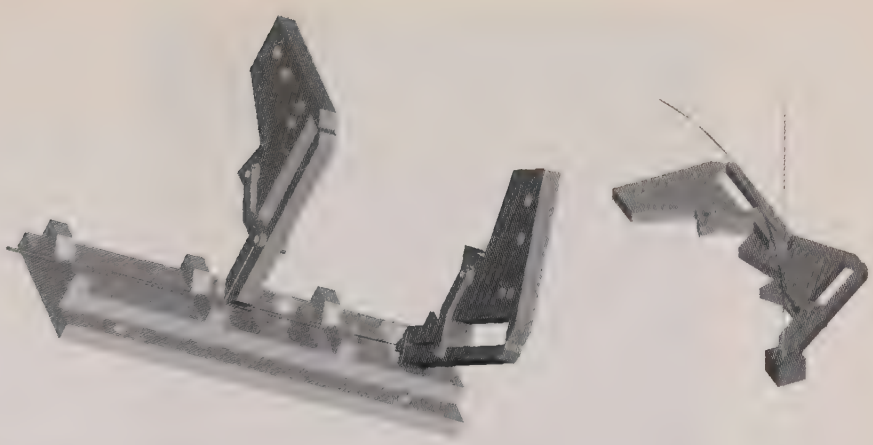


Figure 2.1.34. Gripper and manipulator for fibers [34] (with permission of SPIE).

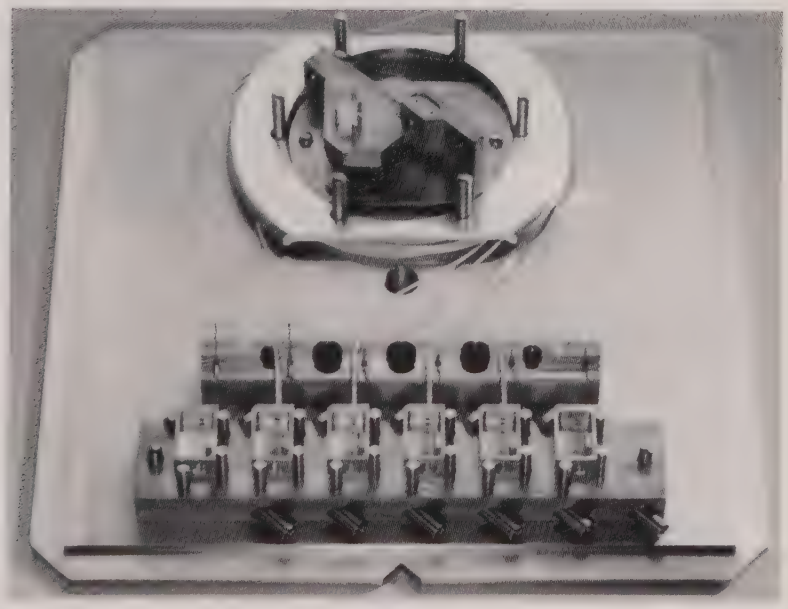


Figure 2.1.35. Work tray containing magazines for material flow during automatic assembly.

position in the overall automatic assembly set-up. The magazines allow the storage of the pieceparts of six duplexer systems, which means that always one tray is in the assembly machine, and a second can be filled off-line with pieceparts in parallel. The diverse grippers and a dispenser for adhesive are mounted on a tool turret, which allows a fast change of tools while occupying only limited space because of its compact construction, as can be seen in Figure 2.1.36.

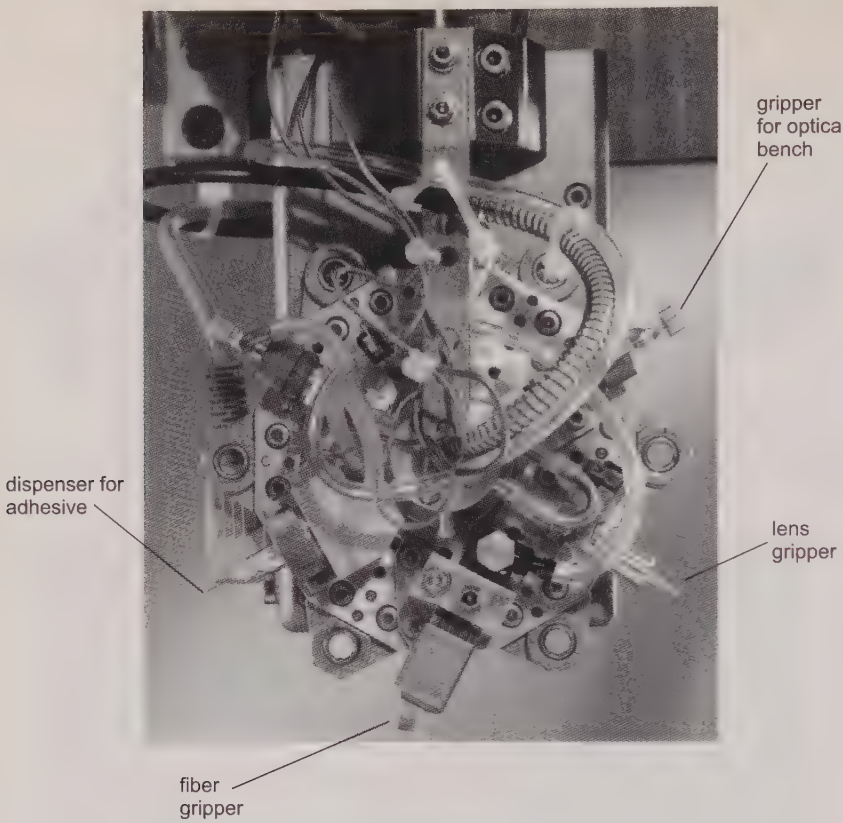


Figure 2.1.36. Tool turret carrying grippers and a dispenser.

One very essential task for alignment and assembly is free sight to the devices no matter whether microscopes are used for manual assembly or cameras for automatic processes. More complex positioning strategies are reported in [38] in the case of flip-chip assembly when the devices are mounted head-over and are hidden one behind the other. As indicated in Figure 2.1.37, a laser chip is to be put on a fiber submount. Both chips carry reference structures as alignment aids. One positioning strategy is explained in Figure 2.1.38. The gripper for the laser chip is also equipped with reference marks. When the chip is attached to the gripper, the relative position of the chip to the gripper can be calculated from the image, which is detected by a camera below the chip. In the next step, the fiber submount is placed under a camera looking from above in order to calculate the relative position of the submount with respect to the gripper. Finally, the flip-chip assembly is done automatically using the calculated coordinates by an *xy*-stage. Another strategy makes use of a prism as an image splitter as shown in Figure 2.1.39. In this case, the camera is placed beside the assembly, a prism is inserted in the free space between the two chips, and the reference marks of both chips are displayed in parallel. This set-up allows a closed-loop control of the alignment.

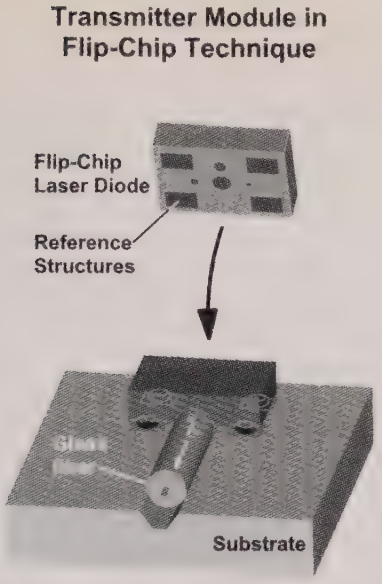


Figure 2.1.37. Flip-chip assembly of a laser chip on a fiber submount [37] (with permission of SPIE).

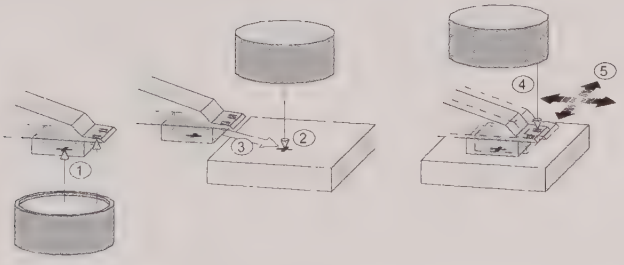


Figure 2.1.38. Alignment strategy for flip-chip assembly using markers on the chip-gripper [37] (with permission of SPIE).

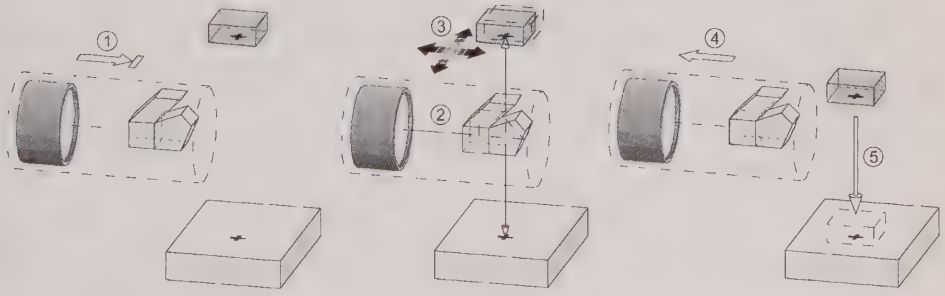


Figure 2.1.39. Alignment strategy for flip-chip assembly using prisms as image splitter [37] (with permission of SPIE).

2.1.7 Conclusions

We have presented a wide variety of modular microoptical systems for telecommunication and sensors. In doing so, we have distinguished between pure opto mechanical devices and opto-electro-mechanical systems. Some of them represent the modular aspect just very basically such as the fiber connectors (Section 2.1.3) or the segmented mirror (Section 2.1.5.1). Other systems follow almost completely the modular method of design and manufacturing such as the distance sensor (Section 2.1.4.3) or the fiber switch array (Section 2.1.5.2).

It is obvious from the wide range of concepts that we have described that the modular approach is convenient and helpful in order to produce more and more complex systems for always higher demands and broader applications. It allows one to fabricate the separate modules with adapted processes, which are not too difficult and time consuming and therefore not too expensive.

On the other hand, one has to ensure that these advantages are not overcome by the costs and problems due to the consequently necessary assembly. Research and development are performed for hybrid solutions, eg, magazines and assembly tools for single components such as lenses or fibers are investigated. However, for a real modular approach to manufacturing and assembly, new strategies need to be found. Microfabricated components should not be first released from their substrate, then be collected in a magazine which perhaps does not provide sufficient position accuracy for a gripper to pick it up again without any need for imaging. In contrast, the method of assembly already needs to be considered in an 'assembly friendly' design of the diverse components and modules and in their fabrication process.

The most important task, however, is the clear and unequivocal definition of interfaces. This is relevant for all aspects of hybrid and modular design and manufacturing, not only for micro-optics but for all kind of systems. It includes interfaces for on-going fabrication such as substrate size and chip size, compatible materials, how an intermediate product is supplied to the next fabrication step, all assembly features including dicing, bonding, and packaging, and – finally and often forgotten – measurement methods with well-defined critical values.

The developers of modular micro-systems have to work hand in hand with the developers of the assembly tools and also of the fabrication equipment. The more they stick to the interfaces and standards, which mostly still need to be defined, the more they can use their developments for several modules or even other purposes. This will make the modular approach economically feasible at least for smaller and medium numbers of pieces such as are often required especially in the sensor field.

2.1.8 References

[1] Deutsche Agenda, *Optische Technologien für das 21. Jahrhundert*, Düsseldorf: VDI-Technologiezentrum, 2000, pp. 3–6.
[2] Stix, G., *Sci. Am.* **284** (2001) 68–73.
[3] Schramm, D., Engel, A., *F&M-Feinwerktechnik* **110**, No. 3 (2002) 34–36.
[4] Kiyoyuki, K., Kikins, R., Jolesz, F.A., in: *Proc MICCAI99*, Vol. 1679 (1999) 1020–1031.
[5] Product Overview, Fiber Optic Connectors, *Europhotonics* June/July (2000) 56–59.
[6] Hibino, Y., Hanawa, F., Nakagome, H., Ishii, M., Takato, N., *J. Lightwave Technol.* **13** (1995) 1728–1735.
[7] Bishop, D.J., Giles, C.R., Das, S.R., *Sci. Am.* **284** (2001) 74–79.
[8] Sawada, R., Higurashi, E., Shimizu, A., Maruno, T., in: *Proc. IEEE/LEOS International Conference on Optical MEMS 2001*, Okinawa, 2001, pp. 23–24.
[9] Marxer, C., de Rooij, N., *J. Lightwave Technol.* **17** (1999) 2–6.
[10] Field, L.A., Burriesci, D.L., Robrish, P.R., Ruby, R.C., *Sens. Actuators A* **53** (1996) 311–315.
[11] Kim, Y., Neikirk, D.P., *Sens. Actuators A* **50** (1996) 141–146.
[12] Debéda, H., v. Freyhold, T., Mohr, J., Wallrabe, U., Wengelink, J., *J. Microelectromech. Syst.* **8** (1999) 258–263.
[13] Mohr, J., Last, A., Hollenbach, U., Oka, T., Wallrabe, U., in: *Proc. IEEE/LEOS International Conference on Optical MEMS 2001*, Okinawa, 2001, pp. 77–78.
[14] Strandmann, C., Baecklund, Y., *Proc. SPIE* **2639** (1995) 18–24.
[15] Hoffmann, M., Voges, E., *Appl. Phys. B* **73** (2001) 629–633.
[16] Wallrabe, U., Dittrich, H., Friedsam, G., Hanemann, Th., Mohr, J., Müller, K., Piotter, V., Ruther, P., Schaller, Th., Zissler, W., *Microsyst. Technol.* **8** (2002) 83–87.
[17] Dunkel, K., Bauer, H.-D., Ehrfeld, W., Hossfeld, J., Weber, L., Hörcher, G., Müller, G., *J. Micromech. Microeng.* **8** (1998) 301–306.
[18] Schulze, J., Ehrfeld, W., Hossfeld, J., Klaus, M., Kufner, M., Kufner, S., Müller, H., Picard, A., *Proc. SPIE* **3737** (1999) 562–571.
[19] Ziegler, P., Wengelink, J., Mohr, J., in: *Proc. 3rd International Conference on Micro Opto Electro Mechanical Systems MOEMS'99*, Mainz, 1999, pp. 186–189.
[20] Krippner, P., Kühner, T., Mohr, J., Saile, V., *Proc. SPIE* **3912** (2000) 141–149.
[21] Hanemann, Th., Heckele, M., Piotter, V., *Polym. News* **25** (2000) 224–229.
[22] Oka, T., Nakajima, H., Shiratsuki, A., Tsugai, M., Wallrabe, U., Hollenbach, U., Krippner, P., Mohr, J., in: *Proc. 11th International Conference on Solid-State Sensors and Actuators, TRANSDUCER'01*, Munich, 2001, pp. 536–539.
[23] Bonner, R., Nossal, R., *Appl. Opt.* **20** (1981) 2097–2107.
[24] Higurashi, E., Sawada, R., Ito, T., in: *Proc. IEEE/LEOS Optical MEMS 2002*, Lugano, IEEE Cat. No. 02EX610, 2002, pp. 33–34.
[25] Peter, Y.-A., Carr, E., Solgaard, O., in: *Proc. IEEE/LEOS Optical MEMS 2002*, Lugano, IEEE Cat. No. 02EX610, 2002, pp. 197–198.
[26] Dellmann, L., Noell, W., Marxer, C., Weible, K., Hoffmann, M., de Rooij, N.F., in: *Proc. 11th International Conference on Solid-State Sensors and Actuators, TRANSDUCER'01*, Munich, 2001, pp. 1332–1335.
[27] Noell, W., Clerc, P.-A., Dellmann, L., Guldimann, B., Herzig, H. P., Manzardo, O., Marxer, C., Weible, K., Dändliker, R., de Rooij, N. F., *IEEE J. Sel. Top. Quantum Electron.* **8** (2002) 148–154.

[28] Göhring, R., Wippermann, F., Kubitz, K., *Proc. SPIE* **4178** (2000) 59–66.
[29] Göhring, R., Wippermann, F., Kubitz, K., Bücker, P., Götz, B., *Proc. SPIE* **3952** (2000) 339–345.
[30] Ruzzu, A., Haller, D., Mohr, J., Wallrabe, U., *Proc. SPIE* **4178** (2000) 67–76.
[31] Kunz, T., Kohl, M., Ruzzu, A., Skrobanek, K., Wallrabe, U., *Microsyst. Technol.* **6** (2000) 121–125.
[32] Krippner, P., Mohr, J., Saile, V., *Proc. SPIE* **3878** (1999) 144.
[33] Krevet, B., Hoffmann, S., Kohl, M., Mohr, J., Oliva, G., in: *Proc. ACTUATOR 2002, 8th International Conference on New Actuators*, Bremen, 2002, pp. 312–315.
[34] Beckert, E., Mohaupt, M., Harnisch, G., Eberhardt, R., *Proc. SPIE* **4755** (2002) 531–548.
[35] Müller, A., Göttert, J., Mohr, J., Rogner, A., *Microsyst. Technol.* **2** (1996) 40–45.
[36] Gengenbach, U., Engelhardt, F., Scharnowell, R., in: *Proc. Symposium on Handling and Assembly of Microparts*, Institute for Precision Engineering, Technical University Vienna, Vienna, 1997, pp. 30–35.
[37] Reinhart, G., Jacob, D., *Proc. SPIE* **4178** (2000) 395–402.

List of Abbreviations

Abbreviation	Explanation
DRIE	deep reactive ion etching
LD	laser diode
MEMS	micro-electro-mechanical system
MOEMS	micro-opto-electro-mechanical system
PD	photodiode
PMMA	poly(methyl methacrylate)
PSD	position-sensitive detector
SEM	scanning electron microscope
LIGA	LIthography, Galvanik (electroplating), Abformung (molding)

2.2 MEMS in RF Filter Applications: Thin-film Bulk Acoustic Wave Technology

R. Aigner, Infineon Technologies, Munich, Germany

Abstract

RF-MEMS filters will replace conventional filters in mobile communication as they offer better performance at lower cost. Requirements and key performance parameters for filters in next-generation mobile applications are reviewed in detail. The fundamentals of RF-MEMS filters – and in particular thin-film bulk acoustic wave (BAW) filters – are summarized in a comprehensive form. The state-of-the-art in BAW filters and future trends are presented. Supplementary information covers manufacturing and commercialization issues.

Keywords: RF-MEMS filter; bulk acoustic wave (BAW) filter; film bulk acoustic resonator (FBAR); solidly mounted resonator (SMR)

Contents

2.2.1 Introduction 176
2.2.1.1 Application Fields for RF Filters 176
2.2.1.2 Types of RF Filters 178
2.2.1.2.1 Ceramic Filters 179
2.2.1.2.2 Surface Acoustic Wave (SAW) Filters 179
2.2.1.2.3 RF-MEMS Filters 180
2.2.1.3 RF Filter Requirements for Mobile Phones 181

2.2.2 Overview on MEMS in RF Applications 182
2.2.2.1 Electrical Field-driven MEMS Resonators 184
2.2.2.2 Bulk Acoustic Wave Resonators 184

2.2.3 Fundamentals of Bulk Acoustic Wave (BAW) Resonators 185
2.2.3.1 Membrane-type BAW Resonators 186

2.2.3.2 Solidly Mounted BAW Resonators 187
2.2.3.3 Material Considerations for the Piezolayer 188
2.2.3.4 Key Performance Parameters for BAW Technology 190
2.2.3.5 Challenges in Processing of BAW Devices 192
2.2.3.6 Diagnostic Methods 193

2.2.4 Bulk Acoustic Wave Filters 194
2.2.4.1 Construction Principle of BAW Filters 195
2.2.4.2 Types of BAW Filters 196
2.2.4.2.1 Ladder Filters 196
2.2.4.2.2 Lattice Filters 198
2.2.4.2.3 Stacked Crystal Filters 199
2.2.4.2.4 Coupled Resonator Filters 200
2.2.4.3 Modeling of BAW Resonators and Filters 200
2.2.4.3.1 Mason Model 201
2.2.4.3.2 Butterworth-van Dyke Model 202
2.2.4.4 Performance of State-of-the-art BAW Filters 203

2.2.5 Commercialization of BAW Technology 206
2.2.5.1 Packaging Requirements 206
2.2.5.2 Reliability Issues 207
2.2.5.3 Future Trends 207

2.2.6 References 208

2.2.1 Introduction

MEMS (micro-electro-mechanical systems) – a technology inspired by microelectronics – has found broad acceptance in the field of automotive sensors and inkjet printheads in the past 10 years. Sensors for pressure, acceleration and mass-flow are today offered by many suppliers. In contrast to this relatively mature market, the application of microsystems/MEMS in the field of mobile communication and high frequency applications is just emerging. The number of mobile phones produced in 2002 is expected to be around 400 million units. The mobile phone manufacturers face ongoing pressure to integrate more functions and features into their products while keeping the prices constant or even lowering them. For this reason they consider microsystems/MEMS to be an attractive solution.

2.2.1.1 Application Fields for RF Filters

Even though there is no widely accepted definition of the frequency span in which a filter will be called an RF filter, in practice the lower limit is set by the maximum frequency at which digital signal processing works reasonably well.

Today this limit is around 100 MHz. The upper limit is around 10 GHz as filters at higher frequencies are usually called microwave filters instead of RF filters. The frequency range from 100 MHz to 6 GHz is most suitable for medium- and short-range terrestrial radio transmission in typical urban environments as the damping is still acceptably low, antenna size is reasonable and sufficient bandwidth can be provided. In industrialized countries all suitable frequency bands are completely occupied by TV broadcasting, mobile phone systems, cordless phones, Bluetooth, wireless local area networks (WLAN) as well as industrial, governmental and military applications. In order to avoid interference between all these applications, high-selectivity RF filters are essential. The selectivity of RF filters defines how much safety margin between adjacent bands will be required.

In terms of market size and unit sales, the RF filters for mobile phones account for 80% of the total market. The number of mobile phones shipped in 2002 is close to 400 million units. All mobile phones need RF filters to protect the sensitive receive path from interference by transmit signals from other users and noise from various RF sources. The minimum receive signal strength at which a phone must still operate can be 120 dB lower (a factor of 10^{12} in power) than the strength of interfering signals. No affordable preamplifier would have sufficient dynamic range and linearity to deal with such a situation. Highly selective RF filters between antenna and preamplifier make sure that only signals from the correct receive band will be amplified. The frequency bands allocated to mobile phone systems can vary from country to country; in general they are in the range 400 MHz–2.2 GHz. The bandwidth is typically 20–75 MHz. The transmit band is located below the receive band but with a gap of only 20 MHz in between. Within this narrow transition range of 20 MHz a receive filter should change from >15 dB attenuation at the upper edge of the transmit band to <3 dB insertion loss at the lower edge of the receive band. In order to achieve such steep filter skirts, the filter elements must have extremely low losses and high quality factors (Q). Selective RF filters are also needed in the transmit path of mobile phones as regulations forbid the emission of RF power outside the specified transmit band. These RF filters take care that the power amplifier will not amplify noise and tones at frequencies outside the transmit band. The European GSM phone system works in a time-multiplexed mode with timeslots reserved for receive and transmit. The antenna in a GSM phone is either connected to the receive or to the transmit path using an RF switch. Owing to this switching the receive and transmit signals are relatively easy to isolate from each other inside a GSM phone. In contrast, the CDMA systems used in USA and the third-generation (UMTS) standard in Europe work in full duplex mode, meaning that the phone receives and transmits simultaneously. This kind of operating mode enforces the use of a so-called antenna-duplexer, which consists of highly selective filters for receive and transmit bands and makes sure that as little as possible power from the power amplifier ends up in the receive path and that the receive signals from the antenna are guided to the preamplifier. A further application of RF filters in mobile phones is channel selection using intermediate frequency (IF) filters in the classical heterodyne receiver. These IF filters typically work at

300 MHz center frequency and have a very small bandwidth. With the advancement of fast AD converters and digital signal processing, most of these IF filters will disappear because channel selection can be done by digital filtering. This approach is called direct conversion or zero IF architecture. The latest generation of GSM phones no longer use IF filters and also in CDMA there is a clear trend towards direct conversion. For all applications together, the number of RF filters used in one mobile phone ranges from 3 to 7, which results in a world market of 2 billion pieces per year.

Another emerging application for RF filters is inside GPS receivers and navigation systems. GPS satellites transmit signals at 1.575 GHz. Signals arriving at Earth are very weak and it is challenging to amplify them in the presence of interfering signals from terrestrial systems. GPS filters require extreme stopband attenuation in order to compensate for the weakness of the signals.

Short range systems such as Bluetooth (2.45 GHz) and cordless phone standards (DECT) also require RF filters. As a contribution to costs, both systems use RF filters with comparably low performance. As a consequence, a greater likelihood of interference with other systems must be accepted. If the number of devices operating simultaneously in close proximity exceeds a certain limit, then the link will not achieve the specified data rate. Emerging applications such as WLAN (5.2 GHz) work close to the upper limit at which short-range transmission is still feasible. Even though this frequency range is less crowded up to now, WLAN has to deal with harmonic tones generated by various mobile phone standards, other wireless applications or even PCs. The RF filters need to have deep notches at certain frequencies.

A large number of remote controls and keyless entry systems work in the so-called ISM bands (industrial, scientific, medical), for example at 433 MHz. These are free bands which are only weakly regulated. RF filters are required in these applications but usually these RF filters do not need to fulfill stringent specifications.

High-end RF filters are used inside basestations for mobile phone systems. The size and cost of these RF filters are less important and therefore they are optimized to have extremely low losses in the passband. The market for these filters is small compared with the market for mobile phones, even though the prices can be 500 times higher. The same is basically true for military and space applications but size and weight advantages can be important in some cases.

2.2.1.2 Types of RF Filters

Practically all the RF filters required for the applications mentioned above are bandpass filters. According to classical filter theory such filters can be realized using inductors (L) and capacitors (C) connected in networks of a certain topology. Depending on the filter order – which is related to the number of reactance elements used – a filter with steep skirts, low insertion loss and good stopband attenuation can be constructed. In practice this approach suffers from the losses

inside inductors and capacitors. At frequencies above 500 MHz even with the best performing inductors it is not possible to built useful filters for the applications mentioned above. The requirements for low insertion loss and steep filter skirts translate into quality factors $Q>400$ if such filters are made out of lumped LC elements. There is no known method to make inductors coming even close to these Q-values at 1 GHz.

2.2.1.2.1 Ceramic Filters

The situation regarding losses is better when using distributed elements such as striplines or electromagnetic cavity resonators instead of lumped LC elements. The wavelength of an electromagnetic wave at 1 GHz in air is 0.3 m. In materials with higher dielectric constant ε_r the wavelength can be reduced by a factor of $\sqrt{\varepsilon_r}$. Ceramic materials with ε_r up to 80 are used for RF filters, materials with even higher dielectric constant tend to have too high intrinsic losses in the frequency range of interest. Ceramic filters use striplines, coupled striplines and other elements with resonant behavior in order to compose the required filter function. Ceramic filters with high selectivity occupy a quite large board space and they are usually the components with the largest height on the printed circuit board of a mobile phone. As the losses are distributed within a fairly large volume they tolerate transmission of several watts of RF power and are robust regarding electrostatic discharge (ESD). The domain of ceramic filters is the antenna-duplexer application in CDMA phones at 1900 MHz because for many years the competing RF filter technologies were not able to fulfill the specifications regarding power handling capabilities and selectivity in this frequency range. Low-performance ceramic filters are also widely used in Bluetooth systems and other short-range applications at higher frequencies. Ceramic filters are delicate to manufacture with sufficient yield as tolerances of the manufacturing processes are critical.

2.2.1.2.2 Surface Acoustic Wave (SAW) Filters

SAW filters dominate the market for RF filters today. Like other acoustic filters they make use of the fact that sound velocity in typical solid materials is factor of 10000 lower than the velocity of an electromagnetic wave. Accordingly the acoustic wavelength is only a few μm at 1 GHz. This enables one to build resonators based on acoustic standing waves with very small size. In order to couple energy between electrical and acoustic domains (and vice versa) the piezoelectric effect in certain crystals is used. SAW filters apply acoustic waves travelling along the surface of a piezoelectric substrate in one lateral direction. So-called Rayleigh waves can be generated and picked up by interdigital combfinger transducers made of metal lines. The wave amplitude quickly decays in the vertical direction so that ideally no energy is emitted in the vertical direction. The lateral period of the interdigital combfingers is designed to fit into the wavepattern of

the Rayleigh waves at the passband frequency. Grating structures between the interdigital transducers can be designed to transmit or reject waves at certain frequencies. Grating type reflectors at both ends of the SAW device take care that acoustic energy stays in the device until finally it will be converted back into the electrical domain. The filter function is defined by the structure of the metal lines in the interdigital transducers, and the gratings. The substrate for SAW filters is usually a monocrystalline piezoelectric material such as quartz, lithium niobate, lithium tantalate and other exotic materials. Manufacturing SAW filters requires a small number of mask steps. This relative simplicity of manufacturing has helped to develop the SAW filter into the most popular filter technology on the market today.

The most severe limitation arises from the fact that the period of combfingers must be very small at frequencies above 2 GHz. This increases the effort and costs for optical lithography dramatically as requirements for linewidth and spacing call for 0.25 μm resolution. SAW filters for frequencies above 2.5 GHz are practically not available today. A second big issue is the power handling capability of SAW filters at high frequencies. In order to transmit 1 W of RF power the narrow combfingers have to carry very large current densities which causes problems like electromigration and overheating. This particular problem disqualifies SAW filters from antenna-duplexer applications in 1900 MHz CDMA phones. The tiny spacing between adjacent combfingers makes SAW filters vulnerable to damage by electrostatic discharge; without additional protection SAWs will not survive even a 0.1 kV ESD pulse. Acoustic stray waves, viscous losses and electrical resistance cannot completely be avoided without expanding the size of the device tremendously. Therefore, the equivalent Q-value of SAW resonators for RF filters is typically lower than 400. Usually the combfingers cannot be passivated and are therefore sensitive to corrosion. This enforces the use of hermetic packages at least for high-performance filters. The best performing piezoelectric substrates such as lithium niobate and lithium tantalate also show pyroelectric behavior, which generates problems in processes using fast temperature ramps. The required piezoelectric substrate is in no way compatible with standard microelectronic manufacturing. Compared with a typical 8 inch raw silicon wafer, the 4 inch substrate for SAW filters is three times more expensive.

2.2.1.2.3 RF-MEMS Filters

Mechanical resonators such as quartz crystals or tuning fork resonators can achieve Q-values well above 10000. Traditional piezoelectric resonators (quartz crystals) are widely used in a frequency range up to 100 MHz. Filters based on tuning fork resonators were widely used for channel filtering in old generation analog wireline telephone systems at frequencies around 100 kHz. Geometrical dimensions define the resonance frequency and therefore MEMS-processes are enabled to expand the concept of mechanical resonators into the GHz range. If the high Q-values can be maintained this enables one to make extremely small filters with excellent performance at low cost. There is no clear distinction be-

tween mechanical and acoustic filters when approaching this frequency range. The principle of RF-MEMS filters will be reviewed in detail in the sections below.

2.2.1.3 RF Filter Requirements for Mobile Phones

Filters (like any other component in an RF system) need to have a specified impedance level at the input and output ports. Many systems work at 50 Ω impedance level because this is a reasonable compromise considering the size and performance of components and the currents and voltages occurring inside the system. In the GHz range, port impedances above 500 Ω become impractical. Filter technologies must be flexible enough to supply RF filters with any impedance between 10 and 500 Ω.

RF signals can either be related to ground (single-ended, unbalanced) or balanced (differential, symmetrical). Balanced components require two connections (in addition to ground) at each port, which carry identical signals with opposite sign. Balanced signals are less vulnerable to interference and crosstalk because interfering signals will only generate common mode amplitudes and thus cancel out in a differential evaluation. A component that converts *bal*anced to *un*balanced signals is called balun. Such a function can be realized using striplines or magnetically coupled inductors, but can also be implemented inside the RF filter. SAW filters preferably perform a balun function by acoustically coupled transducer pairs. While receiver chains benefit from using balanced signals, it is not very common to have balanced antennas; therefore, most receivers need a balun function somewhere between antenna and receiver input. Receive filters featuring this balun function have a growing importance.

The relative width of the frequency bands for mobile phones systems is typically between 2% and 4.3%. The most demanding mobile phone band in this respect is the European GSM1800 band, which operates from 1805 to 1880 MHz (receive) and 1710 to 1785 MHz (transmit). Depending on the type of RF filter, the width of the passband can be challenging to fulfill. A 5% relative bandwidth would definitely be very hard to achieve for acoustic filters unless other requirements are relaxed. Within the passband the insertion loss is one of the most important criteria. State-of-the-art RF filters guarantee a maximum insertion loss of 3–4 dB throughout the whole passband and in a temperature range from –30 to +85 °C. In terms of signal power a 3 dB insertion loss is equivalent to 50% of the power lost. In the receive filter the insertion loss determines the sensitivity of the receiver and the signal-to-noise ratio at a given signal strength. In antenna-duplexer filters a 3 dB insertion loss between power amplifier and antenna also means that a considerable amount of energy is wasted and up to 0.5 W of RF power is converted into heat. Any improvement of insertion loss will be beneficial for the overall performance and efficiency of mobile phones.

The stopband attenuation expected from RF filters depends on the application and is a function of frequency. Between 20 and 30 dB attenuation is typical for

GSM standards, whereas 30 to 40 dB is typical for CDMA standards. At frequencies far above 3 GHz usually less attenuation is needed. Antenna-duplexers for CDMA require up to 50 dB isolation between the receive and transmit ports. A critical issue is always the attenuation that the filter offers at the edges of neighboring bands. Therefore, RF filters should not only have good attenuation far away from the passband, but must also have steep filter skirts. Lower insertion loss and steeper filter skirts both require higher Q-values of the filter elements.

Intermediate frequency (IF) filters are less critical regarding insertion loss because their losses can be equalized by adjusting the amplification of previous stages. IF filters must be very steep in order to allow proper channel selection and their frequency shift with temperature must be very small.

Small size and low cost are extremely important criteria for RF filters to be used in mobile phones. Ceramic antenna-duplexers for CDMA phones are neither small nor cheap and they will disappear quickly because better solutions are emerging [1]. The smallest receive SAW filters on market today have a footprint of 2.0×1.4 mm and a height of 0.6 mm [2]. The average SAW filter price in 2002 is almost four times lower than it was in 2000 and there is ongoing pressure on the prices.

2.2.2 Overview on MEMS in RF Applications

According to the original definition of MEMS, only RF devices with mechanical functionality would qualify as RF-MEMS. As MEMS processes can help to boost the performance of nonmechanical components (such as spiral inductors and other purely electrical components), there is a trend to include them in the MEMS definition also. Table 2.2.1 gives an overview of MEMS devices with possible applications in the RF field and mentions some key advantages.

In addition to RF-MEMS filters – which will be described in detail in the following sections – integrated passive components are a second MEMS-related technology of interest because they save considerable board space and can be included inside multi-chip packages with transceiver ICs in order to make 'system in package' solutions. MEMS processes enable one to fabricate spiral inductors suspended by MEMS structures, which reduces parasitic coupling to the substrate considerably and helps to reduce substrate losses. As the SMD components to be replaced by passive integration are extremely cheap, there is strong pressure to reduce the cost per function. For replacing only RLCs, probably the best choice is to use the cheapest possible substrate and process. On the other hand, the more and more demanding requirements for electrostatic discharge robustness call for complex processes including ESD diodes and other high-performance devices.

A third application field is RF-MEMS switches, which are highly interesting for use in multi-system phones because classical solutions such as PIN diodes or

Table 2.2.1. Application fields for MEMS in RF systems

Category/device	Application examples/MEMS advantage
RF switches (electrostatic)	Band switches, duplex switches, bypass switches; no static power consumption, low loss and signal distortion
RF-MEMS filters and resonators	duplexer filters, RX&TX bandfilters, GPS filters, VCOs, reference oscillators; small size, RF system-on-chip capability
High-Q inductors	'Passive integration', matching components, baluns, VCOs; improved Q-value for on-chip inductors
Tunable capacitors	VCOs, tunable filters; improved Q-value and tuning range for on-chip capacitors

P-HEMT switches have serious drawbacks. Diode switches consume considerable amounts of power for the bias currents in order to be switched on. P-HEMT switches generate distortion and are very sensitive to ESD. Electrostatic actuated MEMS switches can be operated in an extremely power saving mode, they do not generate signal distortion and damage by ESD is less likely. There can also be advantages regarding RF isolation. The main challenges for MEMS in this application are the costs associated with packaging of these components and the reliability over a specified lifetime. Other technical key parameters such as switching speed and required actuation voltage are challenging too, but will be solved. There are both fast (10 μs switching time) and slow switches (10 ms) required in mobile phones. The maximum voltage available in phones is today 2.7 V directly from the battery or up to 10 V using a charge bump. For fast switching devices, the size and mass of the actuator must be minimized. This is only feasible in 'ohmic' switches which use mechanical metal to metal contacts. These contacts are subject to severe degradation following millions of switching cycles and no perfect material for those contacts is available yet. Soft metals which give low resistivity show a tendency for contact sticking and welding. Hard metals need significantly larger contact forces than are available in small MEMS actuators to obtain sufficiently low contact resistance. Considerable research effort will be needed to come up with a satisfactory material combination. In contrast to 'ohmic' switches, the 'capacitive'-type switches consist of moveable plates which are brought into proximity by electrostatic force. There is no DC path through the switch and no metal to metal touchdown. This helps to improve reliability but degrades switching speed because large actuators are needed. The on-state to off-state capacity ratio defines the isolation of the switch. A good isolation calls for large air-gaps in the off-state but this again degrades speed and increases actuation voltage. ESD robustness of capacitive switches is as critical as in P-HEMT switches and additional protection structures are obligatory.

2.2.2.1 Electric Field-driven MEMS Resonators

MEMS resonators at kHz frequencies have been developed, for example, for vibratory gyroscopes [3]. Surface micromachining is the preferred technology as it fits better into a standard IC manufacturing environment. Electric field-driven MEMS resonators are closely related to surface micromachining processes and relatively straightforward to make in this technology. The force F generated by an electric field E is weak even in geometry with small air-gaps. This fact limits the energy exchange between electrical and mechanical domains and the 'coupling'. This coupling is related to the relative frequency spacing between resonance and anti-resonance in energy-coupled systems and in filter applications will determine the width of the passband such a filter can achieve. As F is proportional to E^2 the device is inherently nonlinear and needs a DC bias voltage to work properly. An appealing feature of this principle is the influence of the DC bias on resonance frequency which allows one to make tunable filters. This effect is usually called 'negative electrostatic spring' in the literature. Filters according to this principle have been proposed [4] and the highest frequencies demonstrated are around 100 MHz. In the frequency range below 400 MHz electric field-driven MEMS resonators could be used as IF filters, but unfortunately the next generations of mobile phones will use direct conversion receivers and thus will no longer apply IF filtering. The size of the air-gap is constrained by the electric discharge limit and manufacturing issues. The lateral dimensions of the devices would need to shrink considerably in order to extend the frequency range up to 2 GHz. This would result in an electrical impedance level far above the target range around 50 Ω. As a result of high impedance levels (kΩ range), electric field driven MEMS resonators cannot transmit signals with more than a few mW power because the AC amplitudes will become too large compared with the DC bias voltage. Strong distortion would be the consequence. This prohibits electric field-driven MEMS from use directly at the antenna. In the GHz frequency range the geometric dimensions of the structure will approach the acoustic wavelength and therefore a large number of acoustic modes will become relevant and will potentially cause problems. As the mechanical stiffness of the structure would need to be much higher for a GHz resonator, the benefit of tunability will become less efficient. Serious disadvantages also arise from the fact that air damping in the narrow air-gaps degrades the Q-values dramatically, which means that these filters need expensive vacuum encapsulation.

2.2.2.2 Bulk Acoustic Wave Resonators

This type of MEMS resonator uses piezoelectric materials to drive the mechanical resonance. They benefit from the large driving force F which is linear to the electric field E in piezomaterials. The electromechanical coupling is still a critical parameter, because the stiffness of the resonator is also very high. In the GHz range a mechanical resonator will no longer behave like a 'spring and mass' res-

onator or flexural modes device. Acoustic modes will appear as the resonator size approaches the acoustic wavelength, which in typical solids is in the range 4–11 μm at 1 GHz. A piezoelectric resonator in which a vertical acoustic standing wave is generated within the piezolayer itself is called bulk acoustic wave (BAW) resonator or film bulk acoustic resonator (FBAR). In BAW resonators the resonance frequency is determined by the thickness of the piezolayer, and also by the thickness of electrodes and additional layers in which mechanical energy is stored. In order to avoid acoustic leakage, BAW resonators are either suspended on micromachined membranes [1] or solidly mounted on acoustic reflectors [5]. A detailed comparison of these main types will be given below. As the piezolayer has a relative dielectric constant ranging from about 10 (ZnO or AlN) up to few hundred (PZT), the electric impedance level of BAW devices can meet the standards in RF systems easily. In fact, it is flexible in a range from 10 to several hundred Ω because well behaving BAW resonators can be scaled in area to match the specified impedance. Regarding power handling, BAW devices show excellent performance because linearity is not an issue, no small features that would be prone to failures due to electromigration exist and self-heating can be well controlled. This is in particular true for solidly mounted BAW resonators on Si substrates because there is a very short heat path down to the substrate and the heat sink. BAWs do not require vacuum encapsulation because the vibration amplitudes are extremely small and no squeeze film damping occurs (no tiny air-gaps are present).

2.2.3 Fundamentals of Bulk Acoustic Wave (BAW) Resonators

Acoustic waves in hard solid materials travel with a velocity of typically 5000 m/s and they can travel quite far without decaying (as can be observed if somebody in a far away room drills a hole in a wall). One of the reasons for this effect is that sound waves in solids and in air couple only very weakly to each other as the acoustic impedance of air is factor 100000 lower and thus 99.995% of the energy will be reflected. Quartz crystals which are available up to frequencies of 100 MHz utilize this method to trap the energy inside a thin vibrating plate. The vibrating plate is cut from a monocrystal quartz in a defined orientation and electroded on both faces. A typical 10 MHz quartz crystal has a thickness of 170 μm. The crystal is supported at the edges by springs. The Q-value of such a simple device can be as high as 10000 depending on frequency. The manufacturing of crystals for >100 MHz becomes challenging as the plate thickness approaches 10 μm. In quartz crystals the fundamental mode occurs when $\lambda/2$ is equal to the geometric thickness, the lowest frequency at which a standing wave is possible. The electrodes are very thin compared with the acoustic wavelength and even though no wave effects in the electrodes are relevant the mass loading

shifts the frequency downwards. Quartz crystals can be built for use in thickness shear (TS) mode or thickness extensional (TE) mode or in any mixed mode that is favorable regarding temperature coefficient. The crystal cut determines the mode of operation.

BAW resonators apply MEMS processes in order to extend the principle of quartz crystals to higher frequencies. The typical thickness of the piezolayer is in the range of a few μm or below. Only the TE mode has so far been used because there is no freedom to choose the crystal orientation of deposited piezolayers. The stress field in BAWs looks similar to that in quartz crystals but a larger portion of the standing wave is located inside the electrodes and support layers. This increases the importance of these layers far beyond a slight shift of frequency. The design of the layer stack is an important method for adjusting BAW devices to the specific needs of an application.

2.2.3.1 Membrane-type BAW Resonators (Figure 2.2.1)

The most straightforward way to extend the working principle of a thickness extensional quartz crystal to the GHz range is to build the piezolayer and the electrodes as a membrane structure or on a thin supporting membrane layer. This approach yields a BAW resonator that needs a minimum of layers to be deposited. Forming the membrane can either be done using bulk micromachining processes to remove parts of the substrate or by surface micromachining using a sacrificial layer [1]. As the acoustic impedance of air is a factor of 10^5 lower than in typical solid materials, extremely little energy is radiated into the air at the top and bottom surfaces of the electrodes. The appeal of membrane-type BAWs lies in the small number of layers to be manufactured and in the potentially high Q-values that can be achieved.

On the negative side, it must be stated that layer stress can cause serious problems. As is well known in micromachining, a membrane stack having compressive overall stress will cause severe buckling of the structure. Using membrane-type BAWs implies strict limitations on the freedom to optimize the piezoelectric

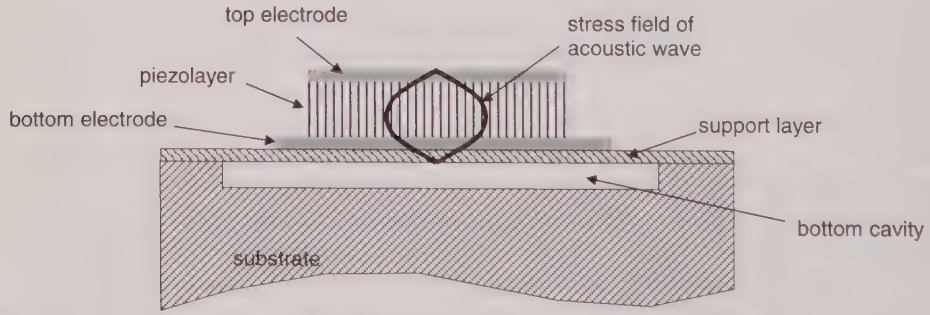


Figure 2.2.1. Schematic cross section of membrane-type BAW resonator.

properties of the layer because the overall stress should stay tensile. Membranes are very delicate to handle as soon as they are released and they are prone to damage during dicing and assembly. Unfortunately, the membrane not only isolates the acoustic waves well from escaping into the substrate but also prevents efficient heat transfer down to the substrate. A large portion of the generated heat will not be removed by convection in air and has to travel along the lateral direction until it finds a proper heat sink. Concerning the power handling capabilities, a membrane-type BAW has some principal drawbacks.

In a membrane-type BAW the designer has to deal with harmonic resonances (overmodes) of considerably high Q-values because the isolation to the substrate is perfect at all frequencies. The second harmonic mode (and higher even modes) can be suppressed by a fully symmetric stack but in this case a third harmonic mode (and higher odd modes) will show up.

2.2.3.2 Solidly Mounted BAW Resonators

As an alternative to membrane BAWs, the acoustic isolation to the substrate can be realized with an acoustic mirror (Figure 2.2.2). Efficient acoustic mirrors can be built using several layers with alternating values of high and low acoustic impedance and a thickness equivalent to a quarter wavelength at the main resonance frequency. This principle of making mirrors is well known from optics where it is used for high-performance dielectric mirrors. At any of the interfaces between high and low impedance layers a large percentage of the wave will be reflected and as the layers are $\lambda/4$ they will sum up with correct phase.

Using three pairs of mirror layers can result in a reflectivity that is good enough for any practical purpose if the impedance ratio $z = Z_1/Z_2$ between the

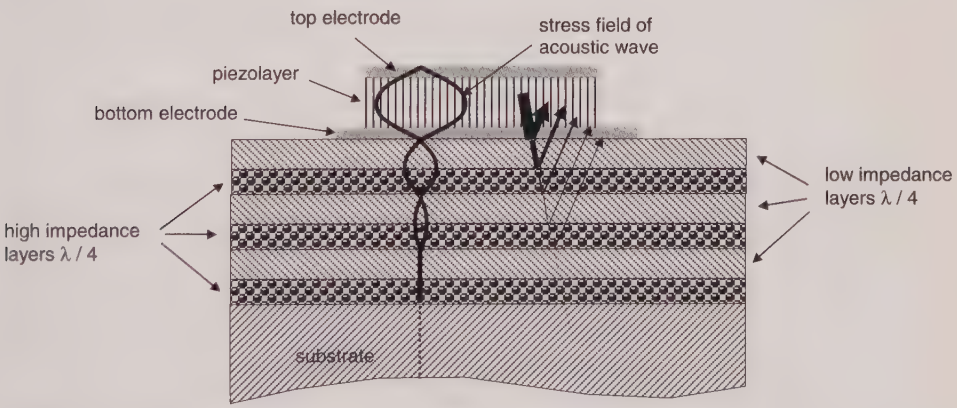


Figure 2.2.2. Schematic cross section of acoustic mirror showing a stress field and the reflections.

layers is high. A general relationship for reflectivity of a mirror with N layer pairs of $\lambda/4$ thickness is

$$r = 1 - z^{2N}$$

Reflectivity is a function of frequency but the plateau of high reflectivity is easily broadband enough to cover the width of typical filter passbands. The behavior of any mirror can be calculated and optimized using the Mason Model as described in Section 2.2.4.3.1 or the transmission line equations known from optics and electrical engineering. The fact that mirror reflectivity is a function of frequency can be utilized to suppress harmonic modes [5]. Harmonic modes are highly damped because the mirror can have bad reflectivity at this frequency.

Manufacturing such a mirror requires several additional layers to be deposited, which increases processing costs. A mirror with excellent impedance ratio z can be made from tungsten and silicon oxide. Even with two pairs an excellent reflectivity can be achieved. A drawback of this mirror is that tungsten is conductive. In order to avoid parasitic coupling between neighboring resonators on a filter chip, the mirror needs to be patterned. If the mirror is made of dielectric layers there is no need to separate resonator areas from each other. The impedance ratios of suitable dielectric mirror pairs is unfortunately limited and typically three pairs are required. Materials with very low impedance such as epoxy or other polymers are not suitable because they have high viscous losses at GHz frequencies. At frequencies below 500 MHz the mirror approach becomes impractical because the $\lambda/4$ layers need to be very thick.

In terms of robustness, such a solidly mounted BAW resonator is superior to a membrane-type BAW. There is no risk of mechanical damage in any of the standard procedures needed in dicing and assembly. There is also no problem with layer stresses in the piezolayer or the electrode layers. For BAWs requiring good power handling capabilities it is very beneficial that a direct vertical heat path through the mirror exists which reduces thermal resistance to the ambient significantly.

2.2.3.3 Material Considerations for the Piezolayer

A precondition for BAWs at GHz frequencies is to use thin-film technology and methods known from IC manufacturing for electrodes and piezolayers. The build-up will start with one of the electrodes, and therefore the piezolayer must grow on top of a metal surface which will not allow epidactical (monocrystalline) growth. Polycrystalline layers can under certain conditions achieve almost the piezoelectric properties of a monocrystal. One of the conditions is a columnar growth and a defined c-axis orientation within the grains [6].

The most popular piezoelectric materials in use for BAW devices are aluminium nitride (AlN), zinc oxide (ZnO) and lead zirconium titanate (PZT). Many more piezoelectric materials are known but most of them are not applicable or

not yet available as thin films. For the performance of BAW devices there are several material parameters which must be considered:

- The piezoelectric coupling coefficient k_t^2. It determines the degree of energy exchange between electrical and mechanical domain. A piezolayer with too low coupling will not be able to make filters with the required bandwidth for mobile phone applications. Having some additional coupling does not hurt and gives some more freedom in filter design. In terms of coupling, PZT is clearly the leader, followed by ZnO and AlN. Details are given in Table 2.2.2.
- Dielectric constant ε_r. The impedance level of a resonator is determined by size of the resonator, by the thickness of the piezolayer and by the dielectric constant. A higher dielectric constant ε_r enables one to reduce the resonator size. AlN and ZnO are similar in this respect with ε_r around 10. This value is clearly beaten by PZT, which can have ε_r up to 400. From acoustic performance considerations a dielectric constant of 100 would be ideal at 1 GHz.
- Sound velocity v_L (longitudinal). A material with low sound velocity will result in thinner piezolayers and thus smaller devices. ZnO and PZT are better than AlN in this respect.
- Intrinsic material losses. Both ZnO and AlN are proven materials in BAW filters. PZT has so far not succeeded in showing sufficiently low intrinsic damping.
- Temperature coefficient. As the piezolayer dominates the resonance frequency it also has a big influence on the temperature drift of the device. AlN has a considerably lower temperature coefficient than ZnO.

The most practicable deposition method for piezoelectric thin films is reactive magnetron sputtering. This works well for AlN and ZnO and these two materials can be reactively sputtered from pure metal targets. AlN is sputtered from an ultra-pure Al target, injecting a mixture of argon and nitrogen gases at low pressures to feed the plasma. The most difficult problem in sputtering of AlN is the high affinity of Al to

Table 2.2.2. Comparison of piezo materials for BAW

	Aluminium nitride AlN	Zinc oxide ZnO	Lead zirconium titanate PZT
Coupling coefficient k_t^2 (%)	6.5	7.5	8–15
Dielectric constant ε_r	9.5	9.2	80–400
Sound velocity v_L (longitudinal) (m/s)	10400	6350	4000–6000
Intrinsic material losses	Very low	Low	High, increase with frequency
CMOS compatible	Yes	No	Never
Deposition rate	High	Medium	Low

oxygen contamination in the chamber. The same method can be used to sputter ZnO; in this case a pure Zn target is used and the plasma is fed by argon and oxygen. Sputtering of PZT is difficult as it requires control of the stoichiometry of lead, zirconium and titanium accurately. In addition to this difficulty PZT belongs to the class of ferroelectric materials (in contrast to AlN and ZnO, which are only piezoelectric), which means that after deposition a poling process is required to align the ferroelectric domains properly. PZT has been successfully applied to MEMS using sol–gel processes, but this approach lacks accuracy in thickness control. A further problem with PZT is the high temperature required to achieve the desired crystallographic phase of the material, because at temperatures of above 500 °C only a few electrode materials such as platinum are still suitable.

Several other material parameters have an indirect effect on the performance of BAW resonators. A high thermal conductivity of the piezomaterial helps to improve power handling capability of a filter. AlN is an excellent heat conductor. Chemical stability can also be an issue for the reliability of devices in a humid environment. ZnO is chemically not very stable whereas AlN is chemically very stable and hard to etch even in the most aggressive acids. Another parameter to be optimized is the breakthrough voltage of the piezolayer. This is related to the bandgap of the dielectric material but also to the defect density of the deposited material.

For industrial applications there are additional issues to be considered in the decision to choose one or another piezo material. Deposition equipment must be mature and reliable. In the likely case that BAWs will be manufactured in a semiconductor environment there are severe contamination issues. Zinc, lead and zirconium are extremely risky materials inside a CMOS fabrication unit as they will severely degrade the carrier lifetime in semiconductor devices. In contrast to ZnO and PZT there is no contamination problem when using AlN. Deposition equipment for AlN is available from several well known vendors of semiconductor equipment, whereas this is not the case for ZnO and certainly not for PZT. Even though AlN seems not to be the ideal material for BAWs at first sight and in theory, it has so far turned out to be the best compromise between performance and manufacturability.

2.2.3.4 Key Performance Parameters for BAW Technology

In a resonator with infinitely thin bottom and top electrodes (and no additional layers above or below), the coupling coefficient of the piezolayer k_t^2 would transform into a relative spacing of series resonance frequency f_s and parallel resonance frequency f_p as determined by the following relationship:

$$k_t^2 = \frac{\pi}{2}\frac{f_s}{f_p}\tan\frac{\pi}{2}\left(\frac{f_p - f_s}{f_p}\right) \cong \frac{\pi^2}{4}\frac{f_p - f_s}{f_p}$$

The spacing of series and parallel resonance will be modified when taking electrode layers and support layers as well as mirror layers into consideration. In most cases

the additional layers will reduce the relative spacing. A method to calculate the series and parallel resonance frequency of an acoustic stack is described in Section 2.2.4.3.1. As series and parallel resonance are easy to measure, an expression equivalent to k_t^2 is introduced. The 'effective coupling' k_{eff}^2 is defined as

$$k_{eff}^2 = \frac{\pi^2}{4} \frac{f_p - f_s}{f_p}$$

k_{eff}^2 is an important parameter for the design of BAW components. Unfortunately, there are various definitions in use by different authors which can differ by as much as 10% from each other [7]. When judging results the definition should be checked.

A second important performance parameter is the quality factor of the resonator. There are several ways to extract the Q-value of a resonator which yield similar results except for pathological cases. The most practical method is to use the maximum of phase steepness as a measure for the Q-value:

$$Q = \frac{f}{2} \left| \frac{\partial \phi_z}{\partial f} \right|_{max}$$

As resonators may have some series resistance or (depending on the piezo material) some shunt conductivity, either the series or the parallel resonance will show the larger Q-value. For this reason it makes sense to define an acoustic Q-value Q_a which is equivalent to the maximum of those two values. In electrical measurements it is simple to distinguish between acoustic losses and electric losses because in a frequency sweep electric losses can be seen even far away from the acoustic resonance frequency where acoustic losses no longer play a role.

For practical applications both a sufficiently high coupling and as large as possible Q-values are the goal. A so-called figure of merit has been introduced to judge the performance of a BAW technology:

$$FOM = k_{eff}^2 Q_a$$

The best published results for FOM are in a range 40–70 [1, 8].

In addition to coupling and Q-values, further parameters are of practical relevance:

- Temperature coefficient of frequency (TCF). Depending on the piezolayer and the other layers used, the resonator will show a temperature drift of frequency. Most available materials cause a negative TCF in BAW resonators; only silicon oxide is known to have a positive TCF. A certain degree of temperature compensation is possible when introducing Si oxide into the layer stack.
- Parasitic substrate effects. Depending on the substrate material used for BAW devices, there will be capacitive and resistive losses down to the substrate which will degrade performance.

- Power handling capabilities. Resonators will heat up when driven at significant power levels. The resonators could overheat if Q-values are too low and insufficient cooling is provided. Therefore, the thermal resistance of a resonator down to substrate is of importance.
- Spurious modes. Good resonators have a clear and smooth impedance response that can be modeled very well with a simple BVD model (Section 2.2.4.3.2). Any deviation of this behavior indicates the existence of strong spurious modes. These modes can degrade the smoothness of the passband significantly, which is not acceptable for certain types of receivers.
- The impedance × area product of a resonator determines how large the BAW component will be. It depends on target frequency, dielectric constant of the material, sound velocity, and the layer stack.

In terms of k_{eff}^2 a membrane-type BAW will theoretically have some advantage compared with a solidly mounted BAW because even a mirror with extreme reflectivity will store some acoustic energy in the uppermost layers. The theoretical advantage in coupling for the membrane BAW is in the range 5–10% depending on the electrode material used. In practice this advantage cannot be utilized because other restrictions apply (layer stress, dispersion, heat conductivity).

2.2.3.5 Challenges in Processing of BAW Devices

Research groups have started and stopped working on BAW devices for as many as 30 years. So far only three companies are able to ship products based on BAW technology [1, 5, 8]. There are numerous challenges and pitfalls in the BAW business, each of which can delay commercial success infinitely.

The most difficult thing to begin with is to achieve the required quality of the piezolayer regarding coupling coefficient. As coupling is linked to correct orientation of the grains during layer deposition, the chances are high that thin films contain too many misoriented grains. It can take years of optimization to achieve coupling coefficients close to bulk values or at least sufficient coupling for some applications. Good and reliable coupling coefficients are a precondition to be able to study all other effects present in BAW devices. Bad coupling usually goes along with bad quality factors. Many severe problems will remain hidden as long as Q-values are below a few hundred. Most likely a prototype BAW resonator will show additional resonances which cannot be explained by one-dimensional theory. These 'spurious modes' can mess up the smoothness of the passband terribly. In the worst case these spurious modes are so strong that extraction of material parameters from electrical measurements becomes impossible. Some of these spurious modes are related to lateral effects in the device and can be improved by a proper design; other types of spurious modes are related to the layer stack itself and require a thorough understanding of the mode shapes that can propagate in the layer stack at relevant frequencies.

Even if prototype BAW resonators show the desired performance, there are more difficult problems to fix. The resonance frequency of a BAW is determined by the thickness of the piezolayer and the neighboring layers. The required tolerance for the resonance frequency is around ±0.1% for typical mobile phone filters, which translates into a thickness tolerance in the same range for the piezolayer and the electrode layers. These extreme thickness tolerances cannot be met by standard tools for semiconductor processes, which typically offer 5% accuracy. Even if the run-to-run variations can be optimized to meet a tighter specification, there is still a major problem regarding thickness uniformity across the wafer to be solved.

2.2.3.6 Diagnostic Methods

The most relevant method for evaluating a BAW resonator is to measure the complex impedance as a function of frequency. This is a standard procedure in RF engineering and can be done using a network analyzer. If this is done in a systematic way using different layer stacks, a set of material parameters of all materials in the stack can be derived including the material coupling k_t^2 of the piezolayer. Material parameters can vary significantly from literature values because deposition conditions may change the Youngs modulus of layers considerably.

Many groups have used x-ray diffraction (XRD) measurements to optimize the crystallographic orientation of the piezolayers. These measurements allow one to judge if the piezolayer qualifies for good coupling or not, but this method is not sufficient to do the final optimization. Even layers with excellent results from XRD may have a bad coupling coefficient because XRD will not reveal the amount of grains with 180° flipped orientation. Each grain with a flipped orientation will neutralize the driving force of one correctly oriented grain and also degrade the Q-value of the resonator. The existence of grains with flipped orientation has been proven [6]. The only realistic method to judge the performance of a piezolayer requires to measure the electrical response of a diagnostic BAW resonator.

Spurious modes can be seen in electrical measurements but their origin is very difficult to localize. The lateral dimension of a BAW resonator will almost certainly be at least 50 times larger than the thickness. This provides plenty of space for lateral modes or plate modes to exist. The theory of acoustic waves in solid materials allows one to calculate the type and amplitude distribution of waves in plates [9, 10]. For a stack of layers with defined thickness which extends infinitely in a lateral direction a so-called dispersion diagram reveals the wavenumber in the lateral direction as a function of frequency. Depending on the frequency there will be one or (at higher frequencies) more branches in such a dispersion diagram, each representing one type of wave that can exist. The wavenumber can either be real, which means that the wave can propagate and carry energy, or imaginary, which represents an 'evanescent' wave that will quickly de-

cay and not carry energy. The dispersion diagram indicates fairly well which modes fit into the lateral dimension of the resonator and can make laterally standing waves at certain frequencies. Knowing the dispersion diagram of the layer stack to be used for the BAW resonator is of fundamental importance. The dispersion diagram can in principle be calculated for any stack (even though this becomes challenging for stacks with many layers), but it requires accurate material parameters including the Poisson's ratio of all layers, which usually is not available. The most efficient method to characterize lateral effects in BAWs and dispersion of layer stacks uses scanning laser interferometry [11]. Complicated optical setups are required to measure the vibration amplitudes of BAWs at GHz frequencies because picometer accuracy is required to identify all relevant modes. An x–y scan of the vibration amplitudes of a BAW resonator at the frequencies of interest is the best method to identify the origin of spurious modes. Which is the best weapon in fighting spurious modes strongly depends on the dispersion type of the active region and the outside region and supporting structure. In some cases a geometry with non-parallel edges is sufficient [1] whereas in other cases this just shifts the spurious modes to different positions without making them less harmful. A highly efficient method has been described [12] that is based on lateral matching of wave amplitudes between the active region and outside region using a boundary region of a certain width and thickness. If applied correctly, very little energy will be transferred to lateral standing waves, thus forcing the resonator to vibrate in a pure thickness extensional mode.

2.2.4 Bulk Acoustic Wave Filters

The electrical response of a BAW resonator with high Q-values is characterized by a large change of complex impedance when sweeping the frequency in a range that includes the main series and parallel resonance (Figure 2.2.3). Far below and far above the resonance the impedance of an ideal resonator is dominated by the static capacity of the resonator and therefore the magnitude of impedance is proportional to $1/f$. When approaching the series resonance the phase will quickly increase from –90° to +90°. The point at which phase passes 0° is defined as series resonance f_s, which is also the point at which magnitude of the impedance is at a minimum. In the frequency range between series and parallel resonance the phase stays at 90° but starts to decrease when approaching parallel resonance f_p where the phase passes through 0° again. The magnitude of the impedance is maximum at f_p. The ratio of the impedance maximum to impedance minimum is approximately equal to the Q-value as long as series resistance of the leads and parasitic shunt conductance are negligible. In general, a good BAW resonator behaves like an almost ideal capacitor below f_s and above f_p and like an almost ideal inductor with varying inductance between f_s and f_p, and it comes far closer to an ideal inductor than any real inductor at GHz frequencies.

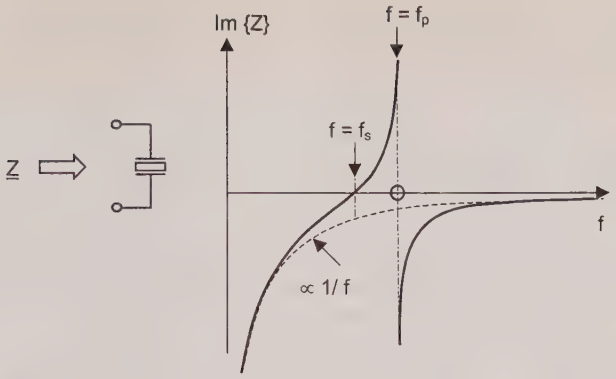


Figure 2.2.3. Impedance characteristic of a BAW resonator.

The classical filter theory which is based on L and C elements is not applicable to BAW filters because the equivalent circuit which represents a BAW resonator (see Section 2.2.4.3.2) can not be 'disassembled' in order to use just the inductance. In terms of filter construction, crystal filters face the same problem.

2.2.4.1 Construction Principle of BAW Filters

BAW filters either consist of several individual BAW resonators which are electrically connected in a certain topology or may also apply acoustically coupled resonators which are stacked on top of each other (this will be discussed in Sections 2.2.4.2.3 and 2.2.4.2.4). In filters which use up to nine individual resonators integrated on one chip it is important to avoid each resonator needing to be tuned to a different resonance frequency for proper operation of the filter because this would mean extremely high processing effort and costs. Typically two groups of resonators having different resonance frequencies will be sufficient to make good filters. The frequency shift between these two groups is realized by an additional layer in the stack which will shift the resonance frequency downwards. This method is called 'detuning' of the resonators. Most BAW filters apply this method. The main design parameter in this case is the size of each resonator which may be varied in a wide range to achieve certain filter characteristics. The designer has to keep in mind that the average size of the resonators must fit the port impedance at which the filter will be used, and not all resonators can be made smaller or larger. The capacity per unit area is determined by the dielectric constant of the piezolayer and the thickness. As the piezolayer thickness decreases as $1/f$ with frequency, the required filter area for standard impedance of 50 Ω will be a function of $1/f^2$. This tremendous size advantage for high-frequency filters may level out at frequencies above 3 GHz because the area required for interconnects will start to dominate. On the other hand, it becomes clear that 50 Ω filters for frequencies below 500 MHz are not very attractive for

making in BAW technology as chips will be large and piezolayers thick, both of which drive the costs up. The typical size of BAW filters will be shown in Section 2.2.4.4.

2.2.4.2 Types of BAW Filters

BAW filters can fulfill all requirements regarding impedance level and port configuration for mobile phone systems. The most straightforward filters to make are single-ended input and single-ended output filters and also filters using balanced signals at both input and output ports (see Section 2.2.1.3). In order to introduce a balun function into the filter, a coupled inductor balun can be integrated in the filter or acoustically coupled 'stacked' structures can be used. The different BAW filter types will be reviewed below.

2.2.4.2.1 Ladder Filters

Ladder filters can in principle be used both for single-ended and balanced signals, but only the single-ended type has practical relevance and will be described here. In ladder filters two groups of resonators are used: series resonators and shunt resonators. The term 'series' is here related to the position of the resonator in the circuit diagram and has nothing to do with f_s. One series resonator and one shunt resonator are connected to a base cell that will be called a 'stage'. Typical ladder-type BAW filters consist of multiple stage or half-stages (Figure 2.2.4), all series resonators at the frequency $f_{s,series}$ and all shunt resonators at the frequency $f_{s,shunt}$, which is shifted down by the detuning frequency f_{detune}.

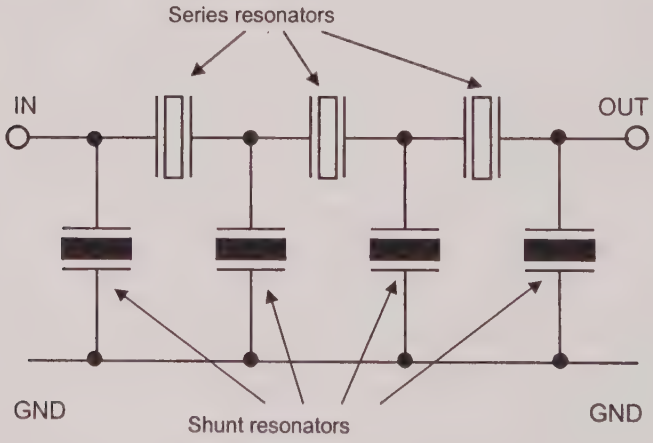


Figure 2.2.4. Topology of 3½-stage ladder filter.

In the center of the passband the series resonators are close to minimum impedance and thus transparent for the signal whereas the shunt resonators are at maximum impedance and therefore will not shunt the signal.

Typical values for f_{detune} are 60–80% of the passband width, which means that $f_{s,series}$ is close to $f_{p,shunt}$. Ladder filters feature steep transitions to the upper and lower stopbands because at the frequencies $f_{s,shunt}$ and $f_{p,series}$ there will be notches with extremely high attenuation owing to the blocking behavior of the series resonator or the shunting of the shunt resonators (Figure 2.2.5).

The far stopbands are dominated by purely capacitive voltage division in each of the stages. The attenuation that can be reached with one stage depends on the capacity ratio of the series and shunt resonators. Large shunts will improve stop-

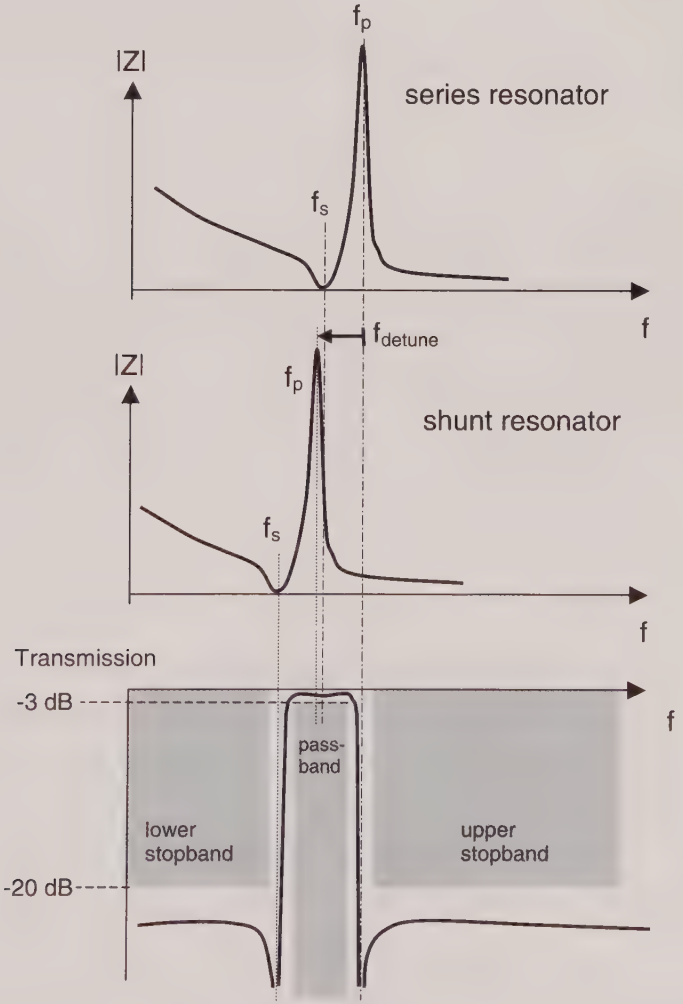


Figure 2.2.5. Working principle of a ladder filter.

bands. There is a certain upper and lower limit for the capacity ratio which is defined by the impedance matching. This becomes more difficult for ratios far from unity. Adding stages helps to improve stopbands but will not result in infinite attenuation while passband insertion loss increases with each stage. Any attempt to increase the filter bandwidth by increasing the detuning frequency f_{detune} will fail because the filter matching in the passband will become badly degraded. The greatest challenge in ladder BAW filters for applications requiring a high relative filter bandwidth (>4%) is to achieve sufficient effective coupling k^2_{eff} in the resonators. For a filter with a 4.3% relative bandwidth the minimum requirement for k^2_{eff} is 6.5% assuming a Q-value of 600. The steepness of the shoulders improves significantly if Q-values can be increased to above 1000, which will also improve the 3 dB bandwidth of the filter as the intersecting points move towards the filter notches. Ladder filters typically have a volcano-shaped passband with a slight buckle in the center. In filters using high-Q resonators the insertion loss is dominated by energy reflected back into the source rather than energy burnt inside the resonators.

2.2.4.2.2 *Lattice Filters*

A filter topology especially suitable for BAW filters is the lattice-type filter (Figure 2.2.6). Unfortunately, this filter type works properly only if both filter ports are balanced (for this reason, a lattice-type BAW filter is often just called a balanced BAW filter). Unbalanced signal components and signals with phase errors will not be filtered very well. In a lattice BAW filter, each stage consists of a bridge-type structure with four resonators. Two series resonators are arranged in diagonal branches while two shunt resonators occupy the other diagonal. As in ladder filters, the series and shunt resonators are detuned by f_{detune}. The working principle of a lattice filter is that either the series or the shunt resonators pass the signal to the output at $f_{\text{s,shunt}}$ or at $f_{\text{s,series}}$, respectively.

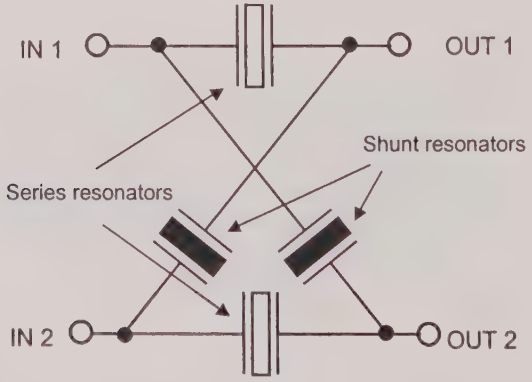


Figure 2.2.6. Topology of a 1-stage lattice filter.

If series and shunt resonators have the same size, the far stopbands will have almost infinite attenuation because the bridge is in perfect balance. Even with one stage an excellent stopband attenuation can be achieved. The width of the passband can for typical cases be up to 30% larger than for ladder filters using resonators with the same for k^2_{eff} and Q-values. The insertion loss of such a lattice filter will also be significantly lower and can be made flatter over the passband. A disadvantage of lattice filters is that they do not automatically generate notches in the transition to the stopband and therefore the filter shoulders are less steep. Introducing notches in a lattice filter is possible by making shunt resonators slightly smaller than series resonators, but this will harm the far stopbands. Using two lattice stages is one solution to achieve good stopbands, and also notches in the transition bands; another method is to combine one lattice stage with one (balanced) ladder stage.

2.2.4.2.3 *Stacked Crystal Filters*

The stacked crystal filter (SCF) is an important component because it allows one to achieve excellent stopband attenuation even in single-ended filters. SCFs are especially suited for very narrowband filters such as GPS filters. For fabrication they require two piezolayers and three electrodes (Figure 2.2.7).

The center electrode will be grounded while the top and bottom electrodes are used for signal input and output. As the center electrode works as an electrical shield, no electrical field can transfer energy from input to output. Only at the mechanical resonance frequency will acoustic energy be transferred from input to output which will constitute the passband. SCFs are more complicated to manufacture because they require two piezolayers with excellent thickness control. As

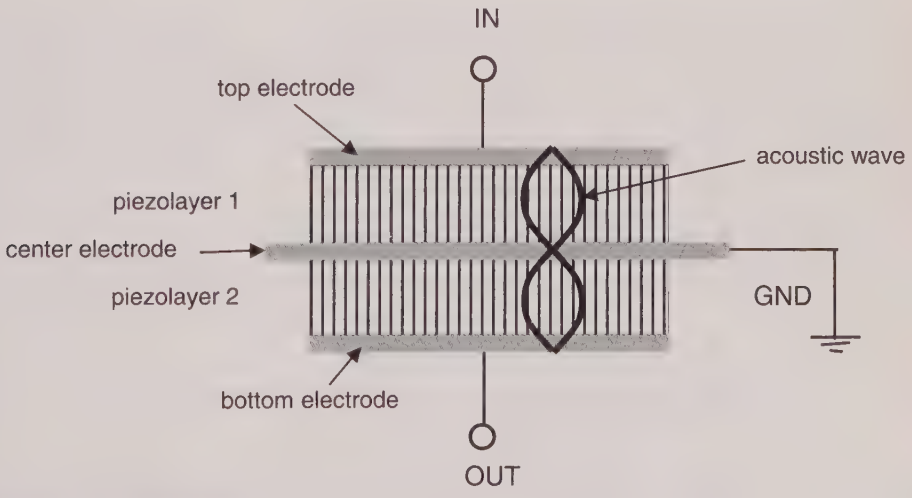


Figure 2.2.7. Schematic structure of an SCF.

there are several wave modes with significant amounts of acoustic coupling, it is beneficial to use SCFs on acoustic mirrors which will damp the unwanted modes [5]. The typical SCF will use an even number of stages because the signals can be fed to the top electrode of the first and last stages while the interstage connections will be buried. No routes through the stack are thus required to connect to the bottom electrode.

2.2.4.2.4 Coupled Resonator Filters

A further extension of the SCF concepts is the recently introduced [13] coupled resonator filter (CRF) concept. This concept has some similarities to piezoelectric transformers used at frequencies of several hundred kHz. Two piezoelectric resonators are acoustically coupled by a specially designed layer stack with a certain acoustic transmissivity in the passband frequency range. CRFs require two piezolayers, four electrodes and a number of coupling layers. Again, it is beneficial to use an acoustic mirror which damps unwanted modes.

The great benefit of CRFs for RF filters is the fact that they are in principle able to perform a balun function (single-ended to balanced signal conversion), even though results on this have not yet been reported in the literature. CRFs should in principle also be able to achieve the same filter bandwidth as a ladder filter using the same piezo material.

Manufacturing of CRFs will be even more challenging than making 'primitive' BAW resonators for the same reasons as mentioned in Section 2.2.3.5.

2.2.4.3 Modeling of BAW Resonators and Filters

Specifications for RF filters in different applications can differ considerably. While some applications rely on very low insertion loss and excellent impedance matching, in other cases the stopband attenuation has top priority. Modeling of BAW devices can be done on different levels. The fundamental physical level would require a three-dimensional coupled electrical and acoustic simulation, which is practically impossible to formulate and solve in an analytical way. Finite element methods (FEM) can in principle be used to solve the problem but are extremely difficult and so far of little practical relevance in BAW development. The two main reasons why FEM is so challenging for BAWs are (i) that a full set of material parameters for all layers is not available and (ii) that the meshing of acoustic structures must be done with a width of $< \lambda/10$ in all dimensions, which can result in element numbers of 10^8. For this reason, even a two-dimensional FEM simulation using 10^6 elements is still very challenging and the agreement with measured results is often disappointing.

The physical level at which BAWs can be modeled efficiently uses one-dimensional acoustic and (piezo)electric equations to describe the stress fields in the layer stack and the electrical impedance at the electrical port of one resonator.

These simulations are extremely important for layer-stack optimization and for the extraction of material parameters. The one-dimensional model of a resonator is called the Mason Model and will be described below. While this type of modeling is most important during development of the BAW resonator, it is too complex for the design of BAW filters and system level simulations. Well behaving BAW resonators can be modeled by a so-called compact (or higher level) model using a simple electrical equivalent circuit known as Butterworth-van Dyke Model.

2.2.4.3.1 Mason Model

The characteristic of a BAW resonator with ideal one-dimensional behavior can be described using a coupled acoustic-electrical model in which all layers are treated as sequence of transmission lines. The piezolayer in addition couples the stress field into electrical voltage and current. The acoustic input impedance Z_{in} of a transmission line can be calculated as function of the frequency f using the following well known equation:

$$Z_{in} = Z_0 \frac{Z_t \cos\left(\frac{2\pi}{v} fh\right) + jZ_0 \sin\left(\frac{2\pi}{v} fh\right)}{Z_0 \cos\left(\frac{2\pi}{v} fh\right) + jZ_t \sin\left(\frac{2\pi}{v} fh\right)}$$

in which Z_0 is the characteristic acoustic impedance and v the sound velocity of the material, h the layer thickness and Z_t the terminating impedance. For multiple layers the above equation can be used subsequently, Z_{in} used as Z_t for the next layer to be added, in order to calculate the acoustic impedance of a stack as a function of frequency. The infinitely large space below the substrate and above the uppermost resonator layer is represented by a termination resistor which has the characteristic acoustic impedance of the material as resistance value. The above equation can be modified to include acoustic losses in each layer but it is tedious work to get any reasonable loss parameters for thin-film material from experiments or from the literature. The impedance load of all the layers above the piezolayer and of all the layers below the piezolayer have a significant effect on the resonance frequency f_s and effective coupling k^2_{eff}. The equation used to include these impedance loads in the electrical impedance characteristic Z seen at the electrical port of the resonator is as follows [14]:

$$Z = \frac{1}{j2\pi fC}\left[1 - K^2 \cdot \frac{\tan \Phi}{\Phi} \cdot \frac{(z_r + z_l)(\cos \Phi)^2 + j \sin(2\Phi)}{(z_r + z_l)\cos(2\Phi) + j(z_r z_l + 1)\sin(2\Phi)}\right]$$

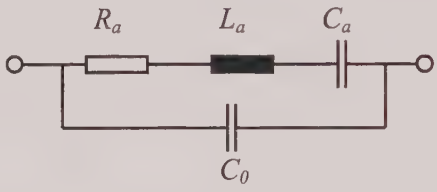
where z_r and z_l denote the upper and lower acoustic load impedance normalized by the characteristic impedance of the piezolayer, $\Phi = f \times \text{thickness} \times 2\pi/\text{velocity}$,

C is the static capacity of the resonator, $K^2 = k_t^2/(1+k_t^2)$ relates to the piezoelectric coupling.

From a plot of this impedance Z over frequency, all relevant resonator parameters can be extracted. This method is suitable for optimizing membrane-type BAWs and also solidly mounted BAWs. For SCFs and CRFs a model equivalent to the approach of Ella et al. [15] is more suitable. In general, the Mason Model gives very reliable results if the material parameters are accurate. The Mason Model is by definition not suitable for modeling spurious modes and other lateral acoustic effects and will also not predict Q-values of resonators accurately. The Mason Model can be extended to include electrical parasites.

2.2.4.3.2 Butterworth-van Dyke Model

The equivalent circuit for a BAW resonator without parasitic effects in the leads is identical with the well known Butterworth-van Dyke Model originally invented for quartz crystals.



The impedance characteristic of the BVD model is practically identical with the results from the Mason Model. The fundamental parameters of a BVD model are as follows:

C	'static' capacity of the resonator	[F]
f_s	series resonance frequenzy	[Hz]
bwr	relative resonator bandwidth	[–]
Q	quality factor of acoustic resonance	[–]

These fundamental parameters can easily extracted from measurements or from simulations using the Mason Model. From impedance measurements it is relatively simple to extract f_s, f_p, and C. Using following equations the fundamental parameters can be estimated:

$$bwr = \frac{f_p - f_s}{f_s}$$

$$C = \frac{-1}{f \times 2\pi \times \mathrm{Im}(Z)}\bigg|_{f<\frac{f_s}{2}} \qquad Q = \frac{f}{2}\left|\frac{\partial \phi_Z}{\partial f}\right|_{\max}$$

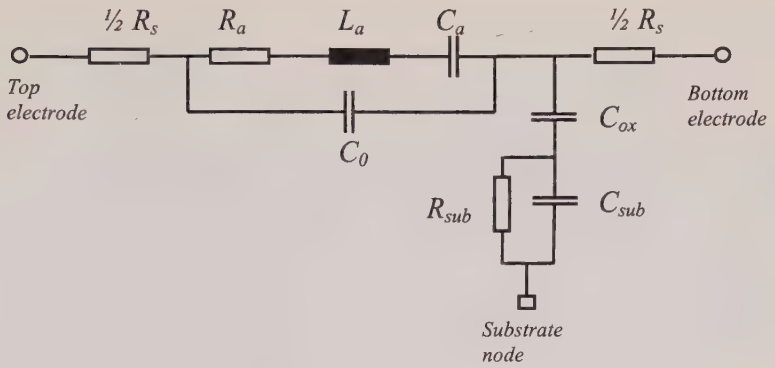


Figure 2.2.8. BVD equivalent circuit including parasitic effects.

From the fundamental parameters, the elements C_0, C_a, L_a, and R_a in the electrical equivalent circuitry can be calculated (Figure 2.2.8). It must be pointed out that they are strictly linked to each other and it is not possible to change any of them individually in an attempt to design a better filter.

$$C_a = C \times 2 \times bwr \qquad C_0 = \frac{C \times 2}{2 + bwr}$$

$$L_a = \frac{1}{(2\pi f_s)^2 C_a} \qquad R_a = \frac{2\pi f_s L_a}{Q}$$

In addition to the fundamental parameters, the dominant parasitic effects can be covered by following additional elements in the equivalent circuit:

R_s	series resistance of one resonator	[Ω]
C_{ox}	capacity between bottom electrode and substrate surface	[F]
C_{sub}	substrate capacity down to chip ground	[F]
R_{sub}	substrate loss resistor	[Ω]

The BVD is a very practical approach to the design of filters and the results will be as close to reality as using any other model. Any circuit simulator will be able to handle a BVD model properly. The BVD can be extended in many ways to include size effects, temperature effects, spurious resonances, and so on.

2.2.4.4 Performance of State-of-the-art BAW Filters

BAW filters are able to beat SAW filters in all performance parameters. The break-even in performance was achieved when an effective coupling of k^2_{eff} of 6.5% and resonator Q of 500 became reality. The advantage over SAW filters at this frequency is an improved insertion loss and steeper filter skirts (see Figures 2.2.10 and 2.2.12).

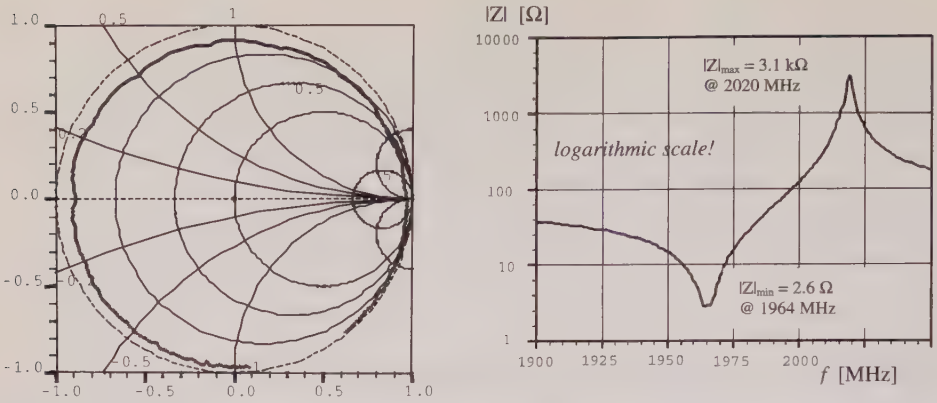


Figure 2.2.9. Smith chart and impedance characteristics of a 1900 MHz BAW resonator. Ratio $|Z|_{max}/|Z|_{min}$ is >1000. Frequency range shown here is 1900 MHz to 2050 MHz.

The use of an acoustic mirror allows one to reduce the temperature coefficient of frequency (TCF) to –22 ppm/K, which is better than SAWs by almost a factor of 2. Chip size is at least a factor of 2 smaller for 1.8 GHz filters. ESD robustness is superior. Power handling up to 3 W is feasible even above 2 GHz, which makes BAW filters an ideal replacement for ceramic filters in duplexer applications.

The figure of merit (Qk^2_{eff}) for the typical resonators that we manufacture is $(900 \times 0.067) = 60$.

Figure 2.2.9 shows a measured Smith chart and impedance characteristics of a 1900 MHz resonator. Figure 2.2.10 shows the simulated influence of resonator

Figure 2.2.10. Influence of resonator Q-values on the performance of a 3½-stage ladder filter (for topology see Figure 2.2.4). Whereas the filter with $Q=400$ barely meets the specification, the filter with $Q=1000$ clearly beats existing solutions.

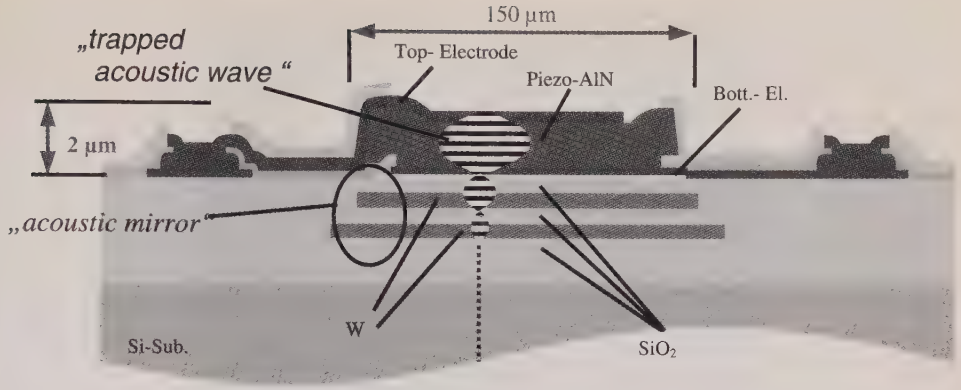


Figure 2.2.11. Schematic cross section of a BAW resonator on an acoustic mirror. Dimensions are given for a 1.8 GHz resonator at 50 Ω impedance level. The vertically trapped acoustic wave is approximately $\lambda/2$ in the piezolayer plus electrodes. The layers of the acoustic mirror have a thickness according to $\lambda/4$. Note: horizontal and vertical directions have different scales.

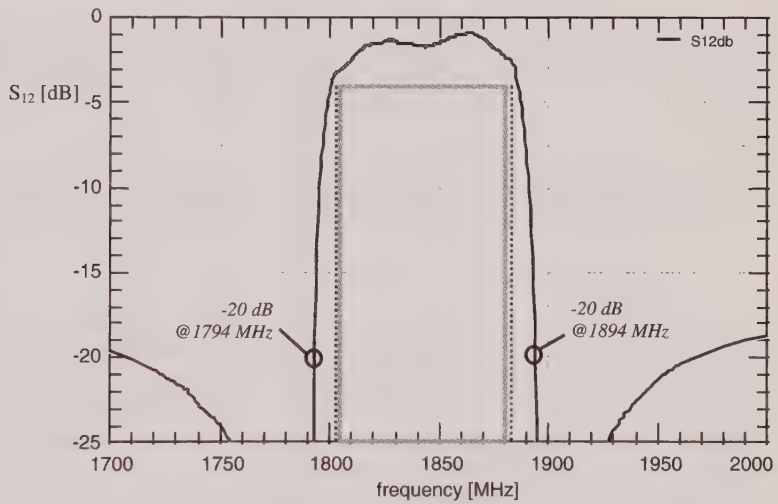


Figure 2.2.12. Measured filter response of a 3½-stage ladder BAW filter. The gray lines are the specification for a DCS/PCN1800 RX filter (passband 1805–1880 MHz). The dotted lines include temperature margins. The 4 dB bandwidth is 86 MHz, the standing wave ratio is <2 throughout the passband. Filter skirts fall from –10 to –40 dB within 5 MHz. Resonator Q-values: 1100. Minimum insertion loss: 0.89 dB.

Q-values on the performance of a DCS/PCN1800 Rx-filter. Figure 2.2.11 shows a cross section of a BAW resonator on an acoustic mirror. The measured response of BAW filters manufactured at Infineon is shown in Figure 2.2.12. A chip photo of a DCS/PCN1800 Rx-BAW filter is shown in Figure 2.2.13.

Figure 2.2.13. Chip photo of a DCS/PCN1800 RX-BAW filter as used in a product from Infineon. Chip size: 0.65×0.95 mm; active area less than 50% of chip size.

The membrane-type BAW technology used by Agilent shows comparable figures of merit [16]. Resonator Q-values are slightly higher while k_{eff}^2 is considerably lower.

2.2.5 Commercialization of BAW Technology

BAW technology has just entered into a battle for market shares against conventional filters and it will take a while until there is a remarkable impact. SAW filters are in maturity, the manufacturers have optimized processes and yields. The cost position of SAWs is hard to beat except for applications in which SAWs work close to their technical limits. The first applications in which BAWs are competitive even at low manufacturing volumes are CDMA-Duplexer filters and other specialty applications. The way to go for BAWs is to lower the manufacturing costs by increasing the production volume to a reasonable level. The next step requires demonstrated benefits at the system level which will increase the expectations of customers regarding the performance and size of filters in general.

2.2.5.1 Packaging Requirements

Packaging costs can be a serious problem in emerging MEMS applications. While mainstream packaging for ICs typically contributes less than 30% to the total manufacturing costs, this ratio can be the inverse for MEMS products. IC

packaging typically uses an overmolding process to form the body of the package. This process cannot be used for most MEMS components because the chip surface must stay clean of any material that could hinder mechanical functionality. This is also true for BAW filters; overmolding would kill the Q-value of the resonators. Compared with other MEMS devices, BAWs are still relatively easy to package because the mechanical robustness is fairly high.

BAWs need a cavity inside the package, and in this respect there is little difference from SAW filters. While SAWs typically need true hermetic packages (in order to prevent humidity from diffusing into the package), it has been demonstrated that BAWs can operate in non-hermetic packages. This advantage will pay off in terms of packaging costs as plastic packages instead of ceramic packages can be used. The final goal for BAWs is to use the 'wafer-level package' approach in which the cavity is provided at waver level [16]. This approach will work a lot better for BAWs than for SAWs because the number of components per wafer is at least a factor of 5 higher.

2.2.5.2 Reliability Issues

BAWs have demonstrated better robustness than SAW filters can achieve with regard to EDS and power handling, even though any new technology and new product need to be carefully evaluated before they are used in high-volume production. The most critical tests include temperature cycling, operation in high humidity and temperature environments, electrostatic discharge tests, repeated reflow soldering, mechanical bending, and drop tests. Improving reliability is a combined effort of chip and package experts.

2.2.5.3 Future Trends

Cost pressure and miniaturization will continue to force the component manufacturers to develop and implement new technologies. The part count in mobile phones will decrease dramatically as modules become more popular. Any module will have to be cheaper than the conventional discrete realization. The mobile phone manufacturers will focus on specifying the functions of modules, integration, and software while the component manufacturers will gather more know-how and responsibility for the physical realization. A clear trend is that as many functions as possible will move into ICs, but it also became clear in recent years that the so-called 'system-on-chip (SOC)' idea is still far away and in commercial terms will probably not make much sense. Integrating all functions into one super-process will increase costs dramatically and will lack flexibility. It is a common understanding today that functions which require specialized technology will be integrated in a hybrid way using the so-called 'system-in-package (SIP)' approach. Finding the best compromise for substrates, assembly, and interconnects regarding performance and costs is a major topic.

BAW technology serves the SIP approach because it simplifies the implementation of RF filters into modules. Performance advantages of the BAW filter may be traded to allow other components to be cheaper. BAW filters in the long term have an impact on system architecture and will grow into a more important component than just a cheaper replacement for SAW or ceramic filters.

2.2.6 References

[1] Ruby, R., Bradley, P., Larson, J., Oshmyansyk, Y., Figueredo, D., presented at the IEEE ISSCC2001 Conference, February 6, 2001.
[2] Meier, H., Baier, T., Riha, G., *IEEE Trans. Microwave Theory Tech.*, **49**, No. 4, (2001) 743–748.
[3] Geen, J.A., presented at the Solid-State Sensor and Actuator Workshop, Hilton Head Island, SC, USA, 8–11 June 1998.
[4] Nguyen, C.T.-C., in: *Proceedings MEMS'98*, Heidelberg, January 25–29, 1998, pp. 1–7.
[5] Lakin, K.M., Belsick, J., McDonald, J.F., McCarron, K.T., presented at the IEEE 2001 Ultrasonics Symposium, October 9, 2001.
[6] Ruffner, J.A., Clem, P.G., Tuttle, B.A., Dimos, D., Gonzales, D.M., *Thin Solid Films*, **354** (1999) 256–261.
[7] Ylilammi, M., Ella, J., Partanen, M., Kaitila, J., *IEEE Trans. Ultrasonics Ferroelectrics Frequency Control*, **49**, No. 4 (2002) 535–539.
[8] Aigner, R., Ella, J., Timme, H.J., Elbrecht, L., Nessler, W., Marksteiner, S., presented at the IEEE 2002 IEDM Conference, San Francisco, December 8–11, 2002.
[9] Auld, B.A., *Acoustic Fields and Waves in Solids*; New York: Wiley, Vols 1 and 2 (1973).
[10] Rosenbaum, J.F., *Bulk Acoustic Wave Theory and Devices*; Boston: Arctech House, 1988.
[11] Graebner, J.E., Safar, H.F., Barber, B., Gammel, P.L., Herbsommen, J., Fetter, L.A., Pastalan, J., Huggins, H.A., Miller, R.E., Optical Mapping of Surface Vibrations on a Thin-Film Resonator Near 2 GHz, paper 3I–4, in: *Proceedings of IEEE 2000 UFFC Ultrasonic Symposium*, Lake Tahoe, USA.
[12] Kaitila, J., Ella, J., Ylilammi, M., *Patent Application WO 2001006647 A1*, 1999.
[13] Lakin, K.M., presented at the IEEE 2002 Ultrasonics Symposium, October 10, 2002.
[14] Lakin, K.M., Kline, G.R., McCarron, K.T., IEEE Trans. Microwave Theory Tech., **41**, No. 12 (1993) 2139–2146.
[15] Ella, J., Tikka, P., Ylilammi, M., Kaitila, J., in: *Proceedings of ECCTD*, 2001, pp. 237–240.
[16] Ruby, R., Barfknecht, A., Han, C., Desai, Y., Geefay, F., Gan, G., Gat, M., Verhoeven, T., presented at the IEEE ISSCC 2002 Conference, February 2002.

List of Symbols and Abbreviations

Symbol	Designation
bwr	relative resonator bandwidth
C	static capacity of resonator
E	electric field
F	force
f	frequency
FOM	figure of merit
f_p	parallel resonance frequency
f_s	series resonance frequency
h	layer thickness
k^2_{eff}	effective coupling constant
k^2_t	piezoelectric coupling coefficient
N	number of layer pairs
Q	quality factor
r	reflectivity
R_s	series resistance
R_{sub}	substrate loss resistor
Z	impedance
z	impedance ratio
ε	dielectric constant
λ	wavelength
v_L	sound velocity

Abbreviation	Explanation
BAW	bulk acoustic wave
CMOS	complementary metal oxide semiconductor
CRF	coupled resonator filter
ESD	electrostatic discharge
FBAR	film bulk acoustic resonator
FEM	finite element method
IC	integrated circuit
IF	intermediate frequency
ISM	industrial, scientific, medical
MEMS	micro-electro-mechanical system
PZT	lead zirconium titanate
RF	radiofrequency
SAW	surface acoustic wave
SCF	stacked crystal filter
SIP	system-in-package
SOC	system-on-chip
TCF	temperature coefficient of frequency

Abbreviation	Explanation
TE	thickness extensional
TS	thickness shear
WLAN	wireless local area network
XRD	x-ray diffraction

PART 3

Sensor Markets

3.1 Automotive Sensor Market*

Karl-Franz Reinhart, Robert Bosch GmbH, Gerlingen
Matthias Illing, Robert Bosch GmbH, Reutlingen

Abstract

Since the late 1970s electronic systems have found increasing use in the automobile. The value of electronics in cars has increased from near zero in the 1970s to almost 20% in a 2000 model mid-class car. In luxury cars this content is even higher. At the same time, electronics are responsible for the majority of innovations and thus contribute strongly to differentiate the car from competitor products. Essentially all of the electronics systems – be it engine management, safety systems or convenience features – need one or multiple sensor signals as input to their signal processing. Over 100 different sensor applications in the car are known; a full-featured luxury car typically contains 100 sensors or more. Therefore, the sensor market has seen at least as strong growth rates as the electronics market over the last two decades and will continue to do so for the foreseeable future. This chapter gives an overview of the automotive sensor market. It starts with a general overview of market size and growth and puts it into perspective with other relevant markets such as the general sensor market and the electronics market. The major distinctive features of the automotive market are explained. Another section details some of the major trends in technologies, products and certain regions. Finally, various market segments are highlighted in more detail. In conclusion, an outlook on possible future developments is given.

Keywords: Automotive; Microsystem; MEMS; Market; Trends

---

* This article is also featured in the forthcoming *Sensors in Automative Technology* edited by J. Marek, H.-P. Trah, Y. Suzuki, I. Yokomori (*Sensors Applications 4*).

Contents

3.1.1	Introduction	214
3.1.2	Automotive Sensor Market Overview	215
3.1.2.1	Market Size	215
3.1.2.2	Technical Issues	216
3.1.2.3	Business Models and Players	217
3.1.3	Incentives and Current Trends in Automotive Sensors	218
3.1.4	Regional Trends	220
3.1.5	Market for Individual Sensors	221
3.1.5.1	Acceleration Sensors	221
3.1.5.2	Pressure Sensors	222
3.1.5.3	Mass-flow Sensors	223
3.1.5.4	Angular-rate Sensors	223
3.1.5.5	Rotational-speed and Phase Sensors	223
3.1.5.6	Angle and Position Sensors	224
3.1.5.7	Exhaust-gas Sensors	224
3.1.6	Outlook: New Sensors and Market Opportunities	225
3.1.6.1	Sensing the Surroundings for Comfort and Safety Systems	225
3.1.6.1.1	Radar and Video Sensors	225
3.1.6.1.2	Night-vision Systems	227
3.1.6.2	New Sensors for Drivetrain and Motor Control	227
3.1.6.2.1	Engine Combustion Sensing	227
3.1.6.2.2.	Sensing of Liquid Media Conditions	227
3.1.6.3	New Sensors for Chassis Systems	228
3.1.6.4	Other Interesting Future Automotive Sensor Markets	228
3.1.6.5	Multiple Sensor Use, Sensor Platforms, and Sensor Fusion	229
3.1.7	Summary	229
3.1.8	References	229

3.1.1 Introduction

According to the association of the German automotive industry (VDA), 55.6 million cars and trucks were manufactured worldwide in 2001 [1]. This figure will not change significantly in the coming years. The automotive market is rather mature – if not saturated – in most countries; average growth rates on the order of just 2% per year can be expected.

However, the picture is different with automotive components. Since the late 1970s, electronic systems have been increasingly used in the automobile. The relative value of electronic components in cars has increased from near zero in the

1970s to almost 20% in a 2000-model-year midclass car and even higher in luxury cars. In addition, electronics are responsible for the majority of innovations and thus contribute strongly to differentiating a car brand from its competitors.

Essentially all electronics systems in use – be it for engine management, in safety systems, or as convenience features – need one or more sensors as input to their signal processing. An overview by Fleming [2] counted 107 different sensor applications in the car; a luxury car typically contains 100 or more sensors. Therefore the automotive sensor market has grown at least as well as the electronics market over the past two decades and will continue to do so for the foreseeable future.

This article gives an overview of the automotive sensor market. It starts with a general overview of market size and growth and puts it into perspective with other relevant markets, such as the general sensor market and the electronics market. The major distinctive features of the automotive market are explained. Another section details some of the major trends in technologies and products and within geographical regions. Finally, various market segments are highlighted in more detail. The conclusion includes an outlook on possible future developments.

3.1.2 Automotive Sensor Market Overview

3.1.2.1 Market Size

Sensors are used in a multitude of markets and applications and perform an innumerable variety of monitoring and control functions. To put the automotive sensor market into perspective, we briefly mention other sensor markets here. According to a study from Intechno Consulting [3], sensor markets can be categorized as follows:

- Machinery manufacturers and suppliers
- Processing industries
- Aircraft and ship building
- Construction sector
- Consumer and other electronics
- Automotive
- Others

According to the same study, in 2003 the overall sensor market will be $42.2 billion, and the automotive sensor market will be about $10.5 billion, or 25% of the overall sensor market. This is by far the largest of the market segments men-

tioned above. Taking into account the tendency of automotive sensors to be lower in cost than sensors in other segments (with the possible exception of the consumer electronics market), the large share of the automotive sensor market is even more pronounced if one looks at production volumes.

The relative importance of the automotive sensor market will continue to grow in the future. The average annual growth rate (CAGR) of the total sensor market from now until 2010 is estimated as 4–5%, and expectations for the automotive sensor market range from 5.1% [4] to 7.5% [3].

It is also interesting to compare the automotive sensor market with the automotive electronics market, since they share the same driving forces and thus have similar dynamics. In comparing market figures, we should be aware that sensors do contain a significant and increasing amount of electronics, so some double counting may occur in the literature.

According to a report of the German association of electrical engineers ZVEI, the automotive electronics market had a volume of $14.9 billion in 2000 and should grow to $23.9 billion in 2005. This is equivalent to a growth rate of 10% per year [5].

3.1.2.2 Technical Issues

As indicated before, about 55.6 million cars and light trucks are manufactured each year. In consequence, each broadly distributed sensor type has a market of about the same magnitude. Due to the high degree of standardization and the platform strategies of original equipment manufacturers (OEMs), typical yearly production volumes for a sensor variant start at about 100000 units and can easily reach several million units per year. In summary, automotive sensors are almost always high-volume products, which sets this market apart from most other sensor markets, except for some consumer applications.

The cost pressure on all automotive components is high. As a result, low-cost, high-volume manufacturing processes are a necessary prerequisite for successfully introducing a sensor into vehicles. This is the reason why certain technologies, such as silicon-based sensors and microsystems, have been particularly successful recently (see Section 3.1.3).

The manufacturing processes need to be reliably controlled to meet high quality standards. To ensure meeting overall automotive quality goals (eg, warranty periods of several years or 100000 miles), the failure rates of sensors must not exceed 10 ppm over their lifetime. While the number of sensors is increasing, these failure limits will be tightened even more. Reliability is especially important in safety-critical applications such as airbags and braking systems. Additionally, in these applications fail-safe or failure-warning mechanisms must be designed into the sensor.

Both the technology and design of sensors thus need to be robust, specially in view of the typically harsh conditions in which sensors operate (Table 3.1.1). To summarize: automotive sensors have to be developed to aerospace specifications and manufactured at consumer prices.

Table 3.1.1. Environmental conditions for automotive sensors

Chemicals and materials	Fuel Oil Water Salt spray Ozone Dust and sand Cleaning materials
Temperature	–40 °C to 125 °C
Vibration and shock	Up to several 100 times gravitational acceleration

3.1.2.3 Business Models and Players

The automotive sensor business is a typical OEM business; there is virtually no consumer aftermarket. The sensors are specified by either the car manufacturer or the tier-one supplier, depending on who has responsibility for a certain system. Car manufacturers and tier-one system suppliers form the customer base for the sensor supplier.

On the supply side, the market can be roughly grouped into three categories:

- System suppliers with in-house sensor component manufacturing
- Semiconductor manufacturers who also have a sensor business
- Sensor specialists

No single sensor supplier offers all – or even nearly all – the 100+ sensor types mentioned above. A few suppliers, however, do offer a broad, comprehensive spectrum of products. Usually these are the system suppliers. Semiconductor suppliers and specialists typically focus on a narrow portfolio, often determined by the technologies or know-how available or necessary for a certain sensor type. As a result, the competitive landscape looks very different, depending on the particular sensor considered.

A (nonexhaustive) list of players is shown in Table 3.1.2.

Table 3.1.2. Nonexhaustive list of major suppliers of sensors to the automotive industry

System suppliers	Denso, Delphi, Siemens, Bosch
Semiconductor manufacturers	Texas Instruments, Motorola, Infineon, Philips, Analog Devices
Sensor specialists	VTI-Hamlin, Sensonor, Murata, Matsushita, Systron-Donner, Hitachi, Kavlico, Nagano Keiki

In terms of business models, system suppliers generally develop and manufacture sensors to support their systems business (captive market). Customized sensors provide a technological and functional advantage to their system. However, due to intense price pressure in the system business, sensors have become an attractive market opportunity on their own. Thus, some systems houses have set up component businesses, sometimes targeting not only the automotive market but nonautomotive markets as well.

For semiconductor specialists, the automotive sensor market provides an extension to their traditional business, with the benefit that it is rather stable and not as cyclical as other semiconductor markets typically are.

The strategy of sensor specialists is to place themselves where specific technologies or application know-how is required.

Generally, the most important success factors in the automotive sensor market are:

- Ability to combine mechanics and electronics using a variety of technologies
- Application know-how
- Long-term strategy: typical development times of 3–5 years or even longer when new technology is needed
- Ability to make significant investments in R&D and manufacturing environment
- Excellent quality management.

3.1.3 Incentives and Current Trends in Automotive Sensors

The applications of automotive sensors can be categorized into the areas of drivetrain, safety, and comfort/convenience. Driving forces for the use of a certain sensor may be market pull (consumer demand), technology push (OEM product differentiation by new feature), or legislation. In the drivetrain and safety areas all three factors are present. However, comfort and convenience applications are driven by market pull and technology push, with legislation usually playing a minor role. Table 3.1.3 gives an overview of applications and drivers for some of the major sensors used in the automobile.

The typical evolution of a sensor over its market lifetime is, first, it is introduced as a technology-push product in a high-end model produced by an OEM. Later, this state-of-the-art technology is recognized as a major improvement in terms of environmental impact or passenger safety and becomes required by law (eg, airbags in the US). Or it is recognized by the customer as an important 'must-have' feature, for example, ESP in Europe. These systems and corresponding sensors are then introduced into all car classes and eventually become a high-volume product.

Table 3.1.3. Driving forces for sensor use in various applications

	Market pull	Technology push	Legislation
Drivetrain	– pressure sensor (direct injection)	– pedal angle (drive-by-wire)	– lambda sensor – pressure – airflow (MAP) – tank pressure
Safety	– speed (ABS) – acceleration (airbag, Europe) – pressure – angular rate (ESP, Europe)	– pressure – angular rate (ESP, USA) – radar/video – surround sensing – acceleration (active suspension)	– acceleration (airbag, USA) – weight sensor (USA) – tire pressure (USA)
Comfort/ convenience	– angular rate (navigation)	– rain sensor (Europe) – light sensor – air quality (HVAC)	

To be able to introduce sensors into all vehicle classes, it is absolutely necessary to manufacture them in high volume at very low cost. This requirement has promoted the development of microsystems technologies, which are often silicon-based technologies. The reason is that silicon technology is a batch technology that offers high volume capability as well as low cost. Developers of such processes benefit from more than 30 years of experience in the manufacturing of integrated circuits, as well as from already having production infrastructure. Sensors that combine electrical and nonelectrical functions in one system are often called microelectromechanical systems (MEMS) or microsystems (MST). Generally, the following trends are supported by these microsystems technologies:

- *Miniaturization.* Typical sizes of functional elements in microsystem-based sensors are on the order of micrometers. A trend to continued shrinkage of feature sizes exists, similar to but not as pronounced as in microelectronics. The overall system size is meanwhile often determined by its packaging and the need for interconnections, rather than by the sensor element itself.
- *Cost reduction.* The requirement to introduce more sensors into all vehicle classes can be met only when the cost of the sensor decreases. A typical example is the acceleration sensor in airbag systems. In the early 1990s an acceleration sensor based on conventional technology cost about $10. Consequently, only high-end cars were equipped with airbags, and side airbags were virtually nonexistent. Introduction of microsystems-based sensors in the mid-1990s offered the same functionality at prices as low as $2. Today, airbags are standard equipment in all vehicle classes.

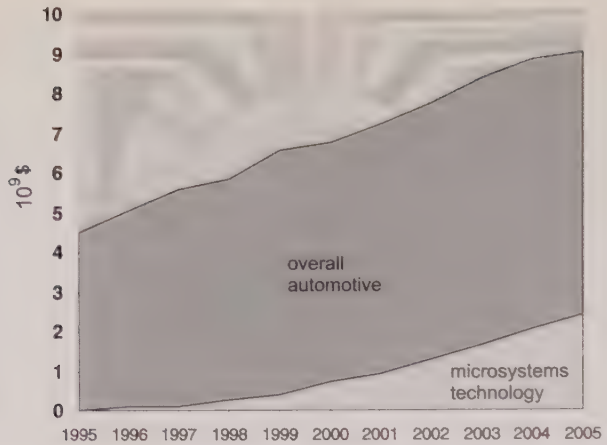


Figure 3.1.1. Market evolution of automotive sensors. Silicon-based sensors are capturing an increasing fraction of the sensor market (adapted from [5]).

- *Increased functionality.* New manufacturing methods offer new opportunities for sensor functionality. Close integration with electronics enables 'smart' sensors; examples include data-preprocessing, continuous self-testing, and digital interfaces. Because of shrinkage of the components, various functionalities can be integrated into one sensor. Commonly combined functions include sensing temperature and pressure or acceleration and angular rate.
- *Quality.* Semiconductor processing techniques enable integration and thus a reduction in the numbers of elements and interconnections, resulting in improved quality. Furthermore, proven batch processes for quality assurance and testing of silicon-based sensors have been adopted from processes for integrated circuits.

The positive features of microsystems technology have led to considerable market success. Although the market for 'traditional' sensors is more or less leveling off, the microsystems sensor market is growing strongly, with predicted growth rates of about 20% per year. By 2005, sensors using these technologies are expected to capture almost 30% of the overall automotive sensor market (Figure 3.1.1).

3.1.4 Regional Trends

Certain regional differences exist. Traditionally, the North American market was a driving force for development of automotive sensors, as well as being the largest market. This trend was due to early US legislation, in particular, for reduction of emission reductions. In consequence, in the US, drivetrain-related sensors play an important role among the sensor types used in the US.

In recent years, the automotive sensor market in Europe has grown strongly and is now of equal importance to the US market. Emission-related legislation has now reached comparable standards as in the US. In addition, the European consumer has demanded safety-related features, many of which, such as ABS and the airbag (including side airbags), are now considered standard equipment.

For the immediate future, several regional trends can be seen. In the US, for example, occupant sensing is an important topic for passive safety systems. Legislation (taking effect in 2003) requires detection of the type of passenger and appropriate modulation of the activation of restraint systems (suppression or low-risk activation of airbags, for example). For this, new sensing systems are required, such as weight sensors, infrared sensors, or video detectors. However, which of these technologies will prevail is still unclear. Other sensors driven by the US market are tire-pressure sensors, for which legislation is currently (2002) being prepared. Another important feature may become rollover detection, using angular rate sensors for passive safety systems, which may be of particular interest for high center-of-gravity sports utility vehicles.

In Europe the ESP has been successfully introduced into the upper and middle vehicle classes. The penetration rate is expected to further increase, so that yaw rate sensors will experience strong growth rates. Another hot topic in the European market is direct-injection diesel (and, more recently, gasoline) engines, which require (among other sensors) high-pressure sensors on the fuel rail and air-flow sensors in the air intake. Another trend is towards higher rates of installation of HVAC systems and new technology within these systems. These will require additional pressure, CO_2, and air-quality sensors.

In Japan, driving forces for sensor use are the high penetration of navigation systems. Another area in which Japan is taking the lead are telematics systems, for which additional sensors will be needed.

3.1.5 Market for Individual Sensors

This section describes the particular markets for major established automotive sensors. Market trends for new sensor types are described in Section 3.1.6.

3.1.5.1 Acceleration Sensors

The use of acceleration sensors has traditionally been driven by passive safety systems, such as airbags and seatbelt tensioners. These systems came onto the market in the early 1980s, with vehicles typically being equipped with one driver-side airbag and a corresponding acceleration sensor. Now, the airbag system in a typical new car includes a front passenger airbag, side airbags on both sides, and window bags and contains 3–5 acceleration sensors.

Although airbags are still the main market for acceleration sensors, other applications require them as well:

- Vehicle dynamics control systems
- Antilock braking systems
- Anti-theft systems
- Active suspension control
- Headlight leveling

The total market for acceleration sensors in 2000, according to NEXUS [7], is $467 million. The market will grow to $687 million by 2005. The main drivers for this growth are the increasing rates of equipment with side airbags. Active suspension control, which typically uses 3–5 sensors per system, has the potential to spur further growth if more vehicles are equipped with these systems in the future.

3.1.5.2 Pressure Sensors

Pressure sensors have numerous applications in the automobile. Historically, the first and still most important is to control the air intake. Other pressure-sensor applications are

- Turbocharger pressure
- Oil pressure
- Atmospheric pressure
- Fuel-tank pressure
- Brake pressure
- Climate control
- Fuel pressure

The pressure ranges and environmental requirements differ from application to application; therefore, various technologies coexist: silicon micromachining and ceramic technologies are mostly used in the low-pressure range (1 kPa to 2 Mpa), and steel membranes are often used for higher pressures.

The market for pressure sensors is still growing, from $700 million in 2000 to $1350 million in 2005, according to a study by Strategy Analytics [4]. The main drivers are fuel-pressure sensors for direct injection engines and higher rates of equipment with advanced air conditioning systems. A strong push for pressure sensors is also coming from tire-pressure monitoring systems, which may well

constitute the largest field of application for pressure sensors by the end of the decade.

3.1.5.3 Mass-flow Sensors

Mass-flow sensors are used to measure air intake. Their advantage over pressure sensors in the same application is primarily higher accuracy. The growth of this sector is moderate and mainly due to increasingly strict environmental legislation.

A wide range of manufacturers use various technologies (heated wire, ceramics, silicon). The market for mass-flow sensors was estimated as about $800 million in 2000.

3.1.5.4 Angular-rate Sensors

Angular-rate sensors are relatively new in automotive applications. They were first used in 1995 in vehicle dynamics control systems. More recent applications include navigation systems and rollover systems. Since all these systems are still in a growth phase, these sensors have a large market potential in coming years. For the year 2000, the market is estimated to be approximately $250 million, growing to more than $700 million in 2005.

3.1.5.5 Rotational-speed and Phase Sensors

Rotational-speed and phase sensors represent the largest market by unit of all sensor types. The most important applications are:

- Wheel speed for ABS and vehicle dynamics control
- Engine speed, camshaft, and crankshaft phase for motor control
- Gear shaft speed for transmission control

Most sensors are based on magnetic field detection. Although previous generations of rotational-speed and phase sensors were passive sensors mainly using variable reluctance principles, current sensors are mostly active, based on the Hall-effect or anisotropic magnetoresistance (AMR), or even (in the future) on giant magnetoresistance (GMR) technology, due to increasing performance requirements.

The total market for rotational-speed and phase sensors is estimated to be 310 million units in 2002, increasing to 390 million units in 2008, and representing a

market value of more than $1600 million. The total market is almost saturated, so new sensor types will replace the older, passive technology.

3.1.5.6 Angle and Position Sensors

Similar to rotational-speed sensors, angle and linear position sensors can be found in a wide variety of applications and systems. Sensor requirements and associated techniques are as manifold as these applications. Some of the most important sensors measure

- Throttle valve angle
- Pedal angle or position
- Seat position
- Gear lever position
- Steering angle

Magnetic principles are used, as well as inductive, capacitive, and optical sensors. In many applications, where cost requirements predominate over performance and reliability issues, simple potentiometers are often found. The market volume in 2002 is about 150 million units and will approach 300 million units, worth more than $1700 million, in 2008.

3.1.5.7 Exhaust-gas Sensors

Pollution legislation continues driving further reduction in engine emissions. The dominant exhaust-gas sensors today and in the near future are oxygen partial pressure sensors – also called lambda sensors. Due to the high temperature of exhaust gas, these sensors are made by ceramics technology in combination with thick-film processing.

Only moderate growth is expected in the market for oxygen sensors, because most cars are already equipped with them. New after-treatment systems for exhaust gas often include more than one oxygen sensor, for example, a lambda point sensor together with a broadband universal oxygen sensor. In the next few years, new concepts for decreasing emissions will increasingly lead to the integration of additional exhaust-gas sensors, such as temperature sensors or partial-pressure sensors sensitive to other exhaust gas components like NO_x or hydrocarbons. Particulate measurements for diesel engines may also result in new types of exhaust-gas sensors, either controlling the function of the particulate filter or being part of a motor component that controls the regeneration of a particulate filter.

Whether use of all of these sensors will become widespread in the near future depends on progress in developing new motor management and injection concepts, which is always struggling to meet the next step of exhaust-gas legislation without making use of expensive additional sensors.

Today's market for exhaust-gas oxygen sensors is about 90 million units, growing to more than 110 million units and to a market value of about $1.5 billion by 2009. The new types of exhaust-gas sensors mentioned above are too new to allow estimating their market penetration in the next generations of systems with any certainty – nevertheless, these new sensor types will finally generate their market.

3.1.6 Outlook: New Sensors and Market Opportunities

Several new sensor types and applications, which can be expected to have significant market potential in the future, are presently in the R&D stage. For most, only qualitative forecasts for market relevance and opportunities can be given, because quantitative estimates in this early development stage suffer from several uncertainties. For example, competing solutions might emerge, or it may turn out that the new parameters can be indirectly derived from existing system data by the use of models. Some of the most interesting new sensor applications in our opinion are mentioned in the following sections.

3.1.6.1 Sensing the Surroundings for Comfort and Safety Systems

3.1.6.1.1 Radar and Video Sensors

Radar and video systems for sensing the surroundings will create an attractive new sensor market in future vehicles and will supplement today's ultrasonic parking aids. The combination and (often) knowledge-based evaluation of various sensor signals as shown in Figure 3.1.2 is a required characteristic for sensing the surroundings of a vehicle. The sensors are not simply detecting a single physical quantity, but, rather, they represent a sensing subsystem aiming to evaluate complex data, as in traffic situations.

A starting point for these new sensor functions is the adaptive cruise control system (ACC), which has already been introduced to the market. ACC systems monitor objects that are far in front of a vehicle and their relative speeds by either RF-radar (at 77 GHz) or laser–radar sensors. Due to the current high price of this system, the market will possibly be restricted mainly to luxury cars in the next few years. Volumes are expected to grow from several hundred thousand units in 2002 to a market of more than two million systems worldwide in 2005 [6].

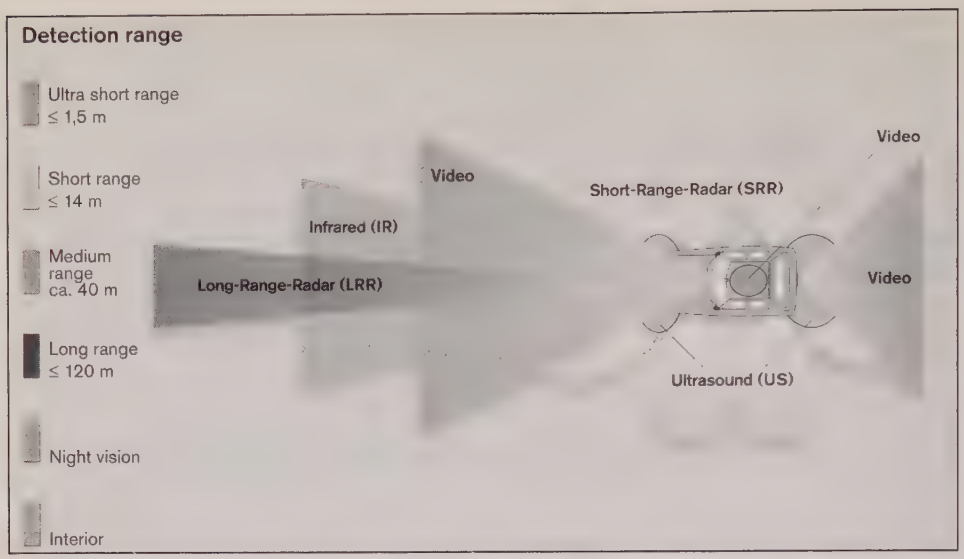


Figure 3.1.2. A combination of sensor technologies can cover the complete surroundings of the vehicle.

Presently, short-range radar systems (SRR) are being developed by several suppliers. Their purpose is to monitor the close environ of the car not covered by ACC systems. Knowledge about nearby objects is required to enhance stop-and-go driving as well as to deal with complex situations in dense city traffic. Presently, short-range detection is done mostly with ultrasonic sensors used to aid parking. Radar sensors at frequencies of 24 GHz will increase the distance range. A system of sensors will allow a new diversity of functions in the fields of passive and active safety, driver support, and vehicle guidance. A significant market for short-range systems is not expected before the year 2005. However, in 2005 about 1 million shipments are expected, representing a considerable market value, due to the high cost of the monolithic microwave ICs (MMICs) to be used in SRR sensors.

Video sensing will develop as a complementary technique for sensing a vehicle's surroundings and can also be integrated with other sensing methods such as radar and ultrasound. Additionally, applications in the passenger compartment, to detect seat occupancy and occupant classification e.g. by height and weight, are possible. Video sensing is encouraged by the considerably decreasing costs of semiconductor video imagers as they become widespread in consumer products and personal computers, as well as by new features of imager chips relating to logarithmic sensitivity and fast random access.

3.1.6.1.2 Night-vision Systems

Night-vision systems have already been introduced in very low volumes in some upscale automobiles. Their function is based on either infrared-sensitive array detectors or a CMOS camera combined with active near-infrared illumination of the scene. The night vision approach further completes the idea of a 'sensitive car'.

3.1.6.2 New Sensors for Drivetrain and Motor Control

3.1.6.2.1 Engine Combustion Sensing

Diesel and gasoline combustion engines can be optimized by using new concepts for controlling combustion in single cylinders. Several sensing principles are under development to monitor combustion, for example, sensing pressure or ion current inside the cylinder, as well as more indirect methods, such as external piezoacoustic sensing. The most precise and direct measurement, however, is achieved by direct pressure sensing in the combustion chamber. Successful development of the corresponding concepts will enable mass production of pressure sensors that are both accurate and able to withstand the extreme environment in a combustion chamber throughout a car's lifetime. This would open a considerable new market, especially considering that one sensor is required for each cylinder. The emergence of combustion-pressure sensors, together with new engine-control concepts, would certainly have a great impact on the market for today's motor-control sensors, because some signals now provided by other sensors could also be obtained from the combustion pressure signal.

3.1.6.2.2 Sensing of Liquid Media Conditions

Another sensor type – the oil-condition sensor – to be integrated into motor management systems is already being introduced into the market. The frequency of routine maintenance tends to increase steadily with each new model year. Making sure that the motor oil remains in good condition for a long time (e.g. 50 000 km) forces the car makers to actually measure the oil's condition rather than estimate it from the miles driven and load data. Oil-condition sensors typically measure physical parameters such as the dielectric constant, viscosity, and temperature to assess the fluid's condition. By 2008, equipment rates as high as 50% can be expected. Engine oil-condition sensors are probably the first members of an emerging new class of fluid-condition sensors for monitoring fluids in general. Applications being investigated include monitoring, for example, fuel type and quality or hydraulic brake fluids. Such sensors could also be used to determine whether conditioners need to be added to an exhaust gas after-treatment reduction system.

3.1.6.3 New Sensors for Chassis Systems

The vision of braking and steering 'by wire' will demand new, extremely reliable sensors. Even in early implementations of steer-by-wire systems, in which manual control can override any system failure, more than one sensor is normally used for the sake of redundancy. Many of the sensor principles required are already established in the market, including steering-angle sensors (eg, for vehicle dynamics control) and pedal-position sensors. Mechanical action or feedback control, however, will drive the emergence of torque and force sensors.

Steering-torque sensors will be needed in new electromechanical steering systems such as active power steering and will therefore have high annual market growth rates in the near future. Several technical solutions are feasible. Torque is measured either directly (eg, by magnetoelastic means or with a strain gauge) or in a more indirect way, by measuring the torque angle of a torsion bar by optical, magnetic, or inductive means. The total market for torque sensors in general can be estimated as approximately 30 million units in 2008. Steering sensors will be the predominant application of torque sensors during the next few years. The market for other torque sensors, such as those intended for drivetrain applications, will be smaller and grow more slowly.

3.1.6.4 Other Interesting Future Automotive Sensor Markets

Several other sensors not described above will enter the market within the next decade, but cannot be discussed here in detail. Just to mention a few:

- Biometric sensors for driver identification (mainly fingerprint for passive enter-and-go systems or for personalized comfort control).
- Gas sensors (eg, for sensing air quality, for measuring carbon dioxide in new types of air conditioning systems, for detecting ozone or fuel-cell gases such as H_2).
- Telemetric sensors (active or passive).
- Electric current sensors (potential-free sensors, eg, for alternators, electric motors, and starters or for battery power management in future 42-V systems).
- Several new sensors might emerge to be integrated into mechanical components, in which the mechanics are part of the signal-generation path but are not produced by the sensor manufacturer. The premise is that business barriers for market entrance, such as product responsibility, or interfacing issues, such as sensor element testing and adjustment procedures, can be removed. Examples include integrated tire sensors or torque and force sensors in gear shafts or on wheel bearings.

3.1.6.5 Multiple Sensor Use, Sensor Platforms, and Sensor Fusion

Presently, automotive sensors are mostly dedicated solely to a particular system. Future sensors will have to be designed as multiple-use sensors for incorporation into different automotive systems. Each sensor then either has to include a bus controller or will be combined – where possible – with other sensors into an intelligent sensor-subsystem that performs a certain level of information preprocessing, for example, an inertial sensor cluster that delivers vehicle dynamics data to the bus. Particularly in the field of inertial sensors will such sensor subsystems emerge in the future – this trend is already being seen in the emergence of multi-axis inertial sensors.

Multiple use of sensors does not necessarily mean that fewer sensor elements are required, because overall sensor performance has to be improved at the same time. The strong demand for fail-safe sensors in safety-related applications will lead to systems using additional sensor signals for increased reliability and based on partial redundancy or on systems to crosscheck the plausibility of the individual signals. Another way of increasing a sensor's accuracy is by using information from multiple sensors.

Sensor fusion and multiple-sensor use will certainly change the long-term sensor market because even more electronics for preprocessing and bus compatibility will be integrated and also because the sensors will become standardized and more easily interchangeable.

3.1.7 Summary

The automotive sensor market is an attractive and still-growing multibillion dollar market characterized by very high production volumes of sensors that must be extremely reliable and low in cost. Microsystem sensor technologies meet these needs and are thus widely accepted in automotive sensor applications. Growth in the sensor market continues to be driven by increasing demands for more safety and comfort and greater reductions in emissions. Existing sensors will continue to find new applications, and new sensors will emerge to improve functionality.

3.1.8 References

[1] Verband der Automobilindustrie Home Page, http://www.vda.de (accessed Oct 2002).
[2] Fleming, W.J. *IEEE Sensors Journal* **1** (2001) 296–308.
[3] Intechno Consulting, Basel, *Sensor Markets 2008*, 1999. See also Intechno Consulting Home Page, http://www.intechnoconsulting.com/ (accessed Oct 2002).

[4] Strategy Analytics, Boston, *Automotive Sensor Demand 1999 –2008*, 2001.
[5] Zentralverband Elektrotechnik und Elektronik Industrie, Frankfurt, *Mikroelektronik-Trendanalyse bis 2005*, 2001.
[6] Strategy Analytics, Boston, *Automotive Radar Systems*, 2001.
[7] Nexus!, *Market Analysis for Microsystems II*, 2002.

List of Abbreviations

Abbreviation	Explanation
ABS	Anti-lock Brake System
ACC	Adaptive Cruise Control
AMR	Anisotropic Magneto-Resistance
CAGR	Compound Average Anual Growth Rate
CMOS	Complementary Metal Oxide Semiconductor
ESP	Electronic Stability Program
GMR	Giant Magneto-Resistance
HVAC	Heating, Ventilation, Air Conditioning
MAP	Manifold Air Pressure
MMIC	Monolithic Microwave Integrated Circuit
OEM	Original Equipment Manufacturer
SRR	Short Range Radar

3.2 Market Data – Sensors in Household Appliances*

Guido R. Tschulena, sgt Sensorberatung, Wehrheim

Abstract

In Europe, the household industry produces millions of washing machines, dishwashers, cookers, refrigerators, home comfort appliances, small appliances, heating and climate control and security and safety-related devices. Here, the use of sensors and microsystem products has a long tradition regarding relatively simple devices, such as temperature sensors or level sensors. However, the rate of introduction of new and improved sensors in household appliances has now increased, and the use of modern sensors with intelligent control systems is a key for differentiation between different appliance products and companies. Therefore, some more general sensor application areas are described, and also specific sensors which have been developed in recent years. Additionally, an outlook is given for possible new products and features which may be introduced in the future.

Keywords: household appliances; sensors; actuators; microsystems; laundry; heating; climate control; market figures; market saturation; European production figures

Contents

3.2.1 Introduction – The Household Appliance Industry 232
3.2.2 HA Industry Data – World Market Data 232
3.2.3 HA Industry Data – European Market Data 233
3.2.4 Sensors and MST in Household Appliances Market Data 237
3.2.5 Conclusion 240
3.2.6 References 241

---

* This contribution is also featured in the forthcoming *Sensors in Household Appliances* edited by G.R. Tschulena and A. Lahrmann.

3.2.1 Introduction – The Household Appliance Industry

The household industry comprises a broad variety of large and small appliances, heating and climate control as well as security and safety devices. The annual turnover figures of the industry are estimated at about 90 billion $ worldwide. East Asian countries represent the largest market, followed closely by Western Europe and the US. Of course, such figures can be nothing more than a rough guide, especially when the sources say nothing about the data, definitions and prerequisites these figures are based on.

The origins of the household appliance industry date back to the early decades of the last century, when "simple" tasks were transferred to household appliances. For example, in an early washing machine of the thirties, the water inlet and outlet as well as water motion and heating were controlled, while all other functions required were carried out manually. Refrigerators only provided the cooling power or the low temperature. In the forties of the last century, the first vacuum cleaners came on the market.

Over the last decades, however, requirements have changed, and the customer expects improved functionality as well as reduced and carefully controlled energy, water and detergent consumption.

Parallel to these customer-driven requirements, we have also seen a change in technology. Electronic control systems and microprocessors in particular have become more powerful, while prices have been falling.

As a result of both developments, the use of automatic control systems has strongly increased. This, in turn, has led to a more widespread use of sensors – not only the traditional low-cost sensors, but a whole host of new sensors that have been introduced – and even more are still being developed.

3.2.2 HA Industry Data – World Market Data

The household appliance field comprises several families of large household appliances, small appliances, heating and climate control devices and security instrumentation. The total turnover of all these appliances is estimated at over 90 billion $ worldwide. The worldwide production of larger appliances could be about 150 million units [1].

The Western European countries provide the biggest market, slightly bigger than the US, followed by the South Eastern countries and Japan, as depicted in Figure 3.2.1.

A moderate growth of these production figures can be expected – depending on many socio-economic and technical factors. So the European appliance production is expected to grow by about 1 to 3% to about 65 million units in the next five years.

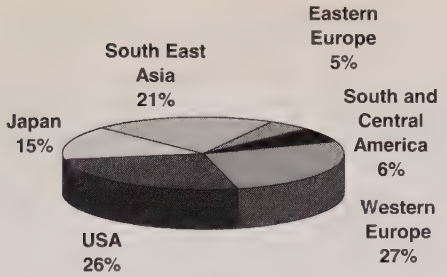


Figure 3.2.1. Geographical distribution of the large and small appliance markets according to production quantity.

3.2.3 HA Industry Data – European Market Data

The European consumption of appliances is in the range of several million units per year, for example up to 50 million units per year for washing machines or refrigerators, as shown in Figure 3.2.2 [2].

European production can only partly cover these needs, while the rest is imported. Nearly 70 million units of large appliances per year are produced in Europe.

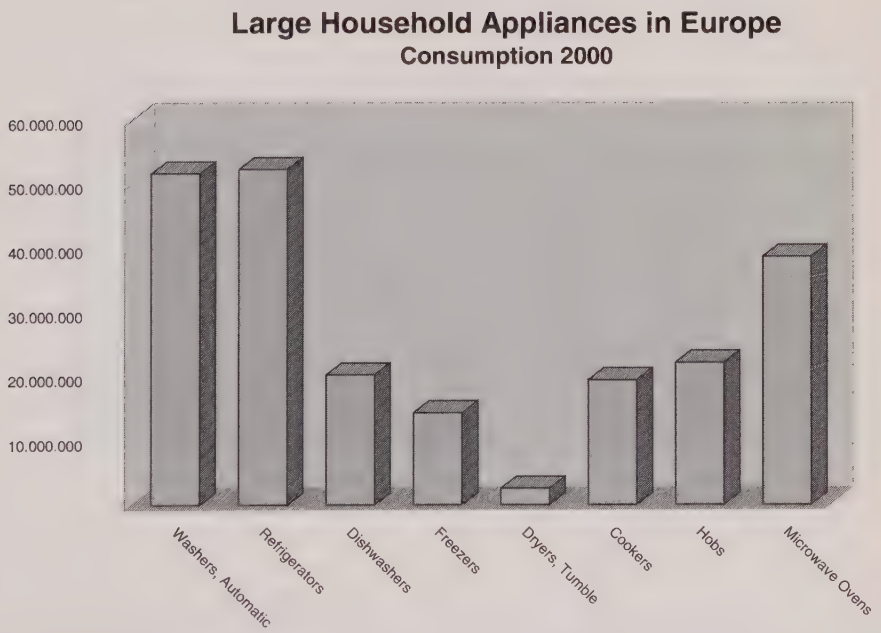


Figure 3.2.2. European consumption of selected larger household appliances.

The turnover of the household appliance industry is about 35 billion Euro, of which about 63% come from large appliances, 17% from small appliances, and 20% from heating, ventilation and air conditioning appliances.

A workforce of about 200,000 are directly employed by the European household appliance industry, and about a further 500,000 employees depend on it, working in upstream or downstream jobs [3].

The European production figures for large appliances are given in Figure 3.2.3 [2].

- Automatic washing machines about 20 million units
- Refrigerators about 18 million units
- Microwave ovens, dishwashers and cookers, each between 6 and 7 million units
- Built-in ovens, and dryers, each somewhat more than 3 million units

European production figures for selected larger appliances.

The largest share of appliances in Europe is produced (measured in terms of quantities) in Italy (with about 38%), Germany (about 23%), the UK, France, Spain, Turkey (about 7%), followed by Russia and Slovenia (about 2%).

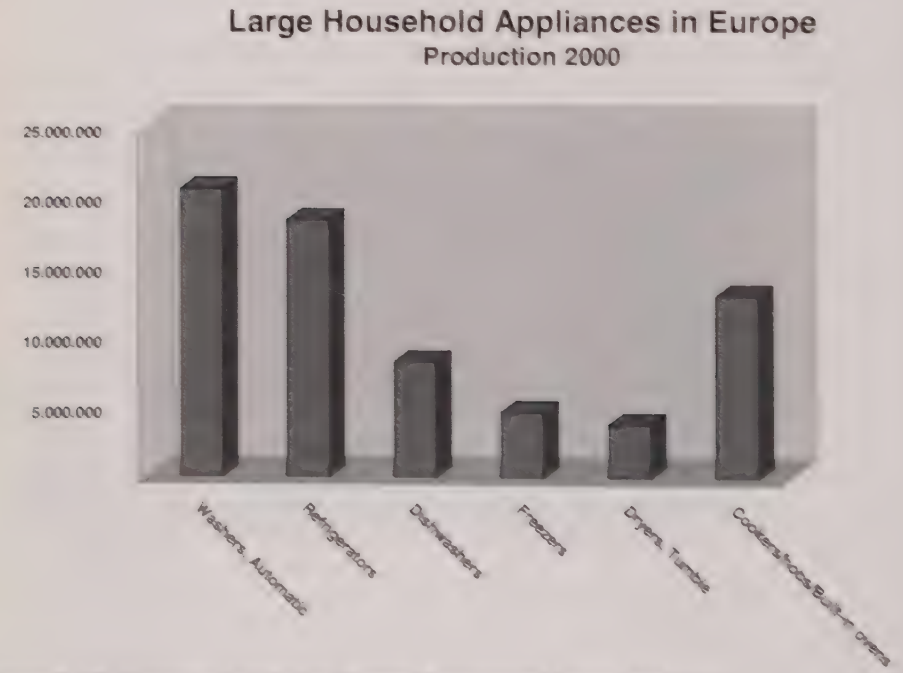


Figure 3.2.3. European production figures for selected larger appliances.

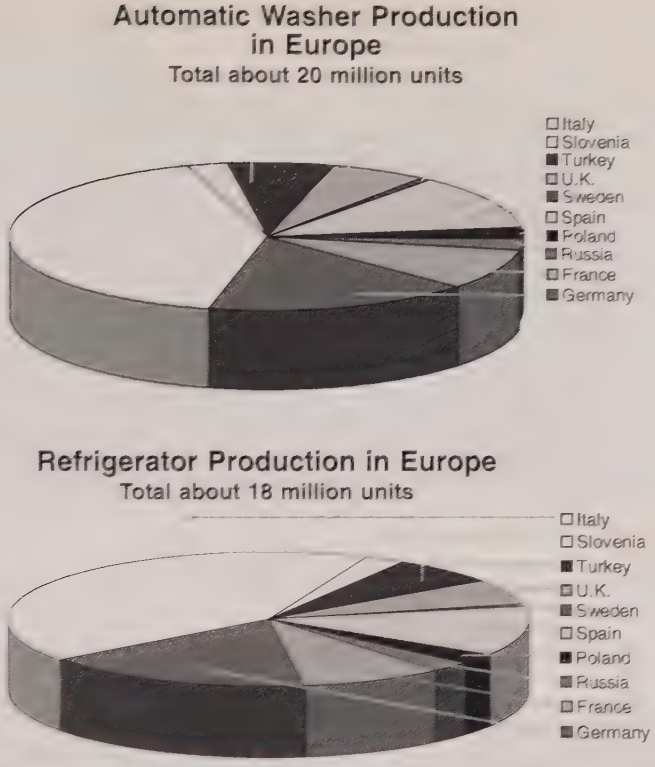


Figure 3.2.4. Geographical distribution of the appliance production in Europe for washing machines (top) and for refrigerators (bottom).

A moderate growth of these appliance production figures can be expected – depending on many socio-economic and technical factors as well as entrepreneurial activities among competing companies and brands.

The degree to which European households are equipped with appliances varies – the saturation level being highest in Central and Northern Europe, somewhat lower in Southern Europe and still lower in Eastern Europe. The average saturation level for some larger appliance families is shown in Figure 3.2.5. For washing machines, cookers and refrigerators, a saturation of more than 90% has been achieved, around 50% for microwave ovens or freezers, and only 25% for dryers.

After a new product has been introduced, market behavior develops over time, and a higher degree of saturation will eventually be reached. The data given in Figure 3.2.6 refer to Germany. The eighties saw the introduction of microwave ovens, whose market share has been steadily increasing to the present level of over 60%. By contrast, in the case of washing machines and refrigerators, which are found practically in every household, the saturation level remains more or less constant, as business consists mainly in replacing existing appliances plus a small share of new households.

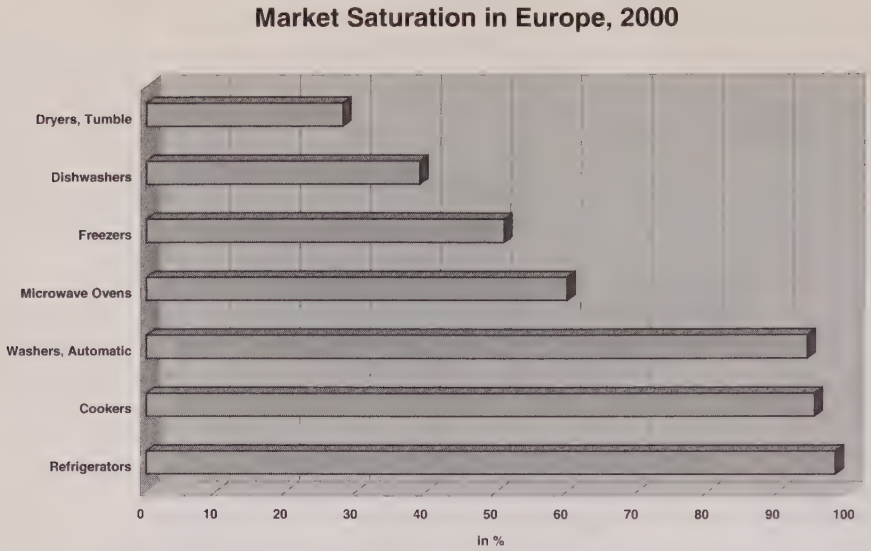


Figure 3.2.5. Market saturation of selected large appliances in Europe for the year 2000.

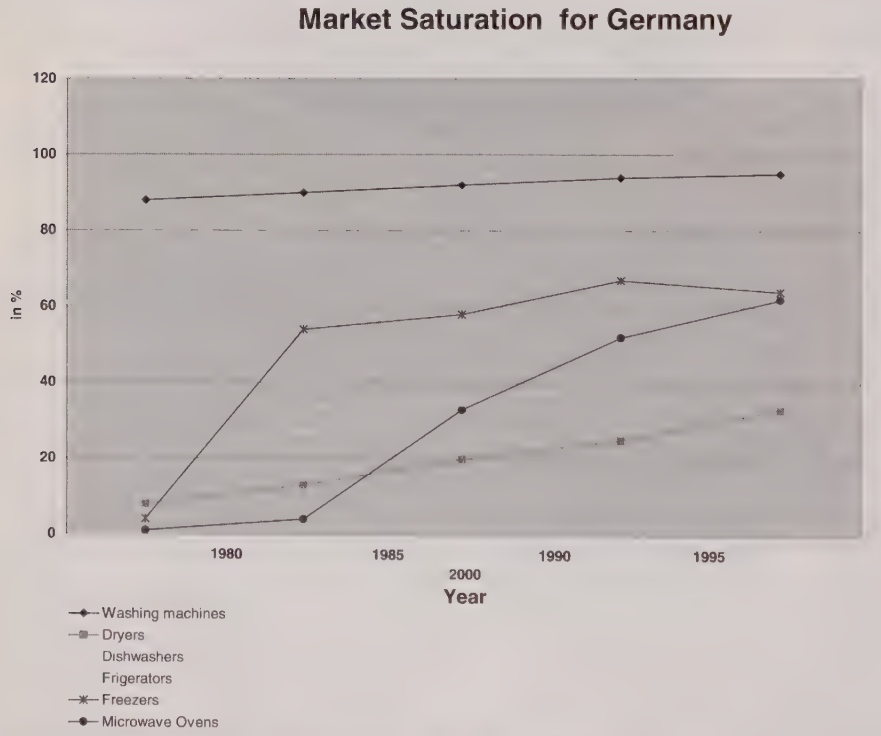


Figure 3.2.6. Development of the market saturation for selected appliances [4]. Figures refer to Western Germany before 1990, and reunited Germany from 1990.

The appliance industry in Europe comprises about 280 companies, of which 90% are small and medium-sized independent enterprises. The larger companies include Electrolux (with a market share of approximately 20%), Bosch-Siemens Hausgeräte (about 16%) and Whirlpool (about 11%). These larger groups are followed by Merloni (Italy), Groupe Brandt, General Electric, Candy, Miele (Germany) and Arcelic (Turkey), and many more smaller appliance companies.

There is only limited scope for growth in the European household appliance industry, as could be shown for large appliances. While quality has been improved – e.g. improved energy efficiency – average real prices of the appliances have continued to drop over the past years. As a consequence, pressure to cut costs has been building up and led to the transfer of production facilities to countries with lower labor costs 8Stamminger, 2002).

With the introduction of more sophisticated, consumer-friendly appliances, this trend could be overcome, and European competitiveness could be strengthened. The use of better control systems and the introduction of more and better sensors are a prerequisite for such a development, as we are going to point out in this book.

3.2.4 Sensors and MST in Household Appliances Market Data

The function of an appliance determines the choice of sensors. Thus, even in the early appliances, e.g. in the forties and fifties of the last century some relatively simple sensors were used.

The most common control functions in these early appliances are the control of temperature, pressure position or distance. Mechanical sensing devices were introduced for these purposes, such as bimetal temperature switches or liquid expansion temperature switches for ovens, washing machines, dishwashers, refrigerators, etc. Electromechanical pressure switches and potentiometric level sensors have also been introduced quite early.

When microelectronics and solid state devices developed over the last five to four decades, the development of solid-state sensors followed suit, resulting in the introduction of NTC and PTC resistors to monitor temperature, and first Reed relais and inductive sensors to determine position and distance, or tachometers for rotational measurements in washing machines and dish washers over the past two decades.

An overview of sensors in selected appliances is given in Table 3.2.1.

The sensor industry is particularly strong in Germany and the rest of Europe – about 2000 companies in Europe are producing or trading sensors, more than 600 companies in Germany, which are mostly highly specialized traditional small and medium-sized enterprises. They have developed a wide range of new devices which they sell to various industries, such as the process, manufacturing and ma-

Table 3.2.1. Examples of sensors in appliances used today [5, 6].

Appliance	Monitored function	Type of sensors used today
Washers	Temperature	NTC, Klixon
	Rotation	Tachometers
	Level	Pressure sensors
	Weight	Inductive sensors
	Position	Reed switches
	Unbalance	Tachometer, pressure sensors
Dryers	Temperature	NTC, Klixon
	Humidity	Conductivity sensing
Dish washer	Temperature	NTC
	Salt level	Reed contact + swimmer Density sensors
	Cleanser level	Reed contact
	Spray arm rotation	Magnetoresistive sensors
	Flow	Rotary water meters Pressure sensors
Freezers	Temperature	
Oven	Temperature	Expansion capillars, Pt 100
Microwave ovens	Temperature	NTC
	Humidity	Ceramic sensors
	Gases	
Hair dryers	Temperature	NTC
	Temperature (contactless)	IR Thermopiles
	Humidity	
	Air flow	
Toasters	Temperature	NTC
	Temperature (contactless)	IR Thermopiles
Irons	Temperature	Bimetal switch, NTC
	Humidity	
	Position, movements	Tilt sensors, acceleration sensors
Vacuum cleaners	Air flow	
	Particles	
Air cleaners	Gases	Ceramic sensors
	Particles	

chine building, automotive, environmental protection industries. Some of them, however, are exclusive suppliers to the household appliance industry.

The industry has its technological roots in the precision mechanics technology, in particular in Switzerland, in Germany and in some French regions (the former watch-making regions).

Over the past two or three decades, new technologies have been developed and introduced, often based on semiconductor technology. The trend in micro-

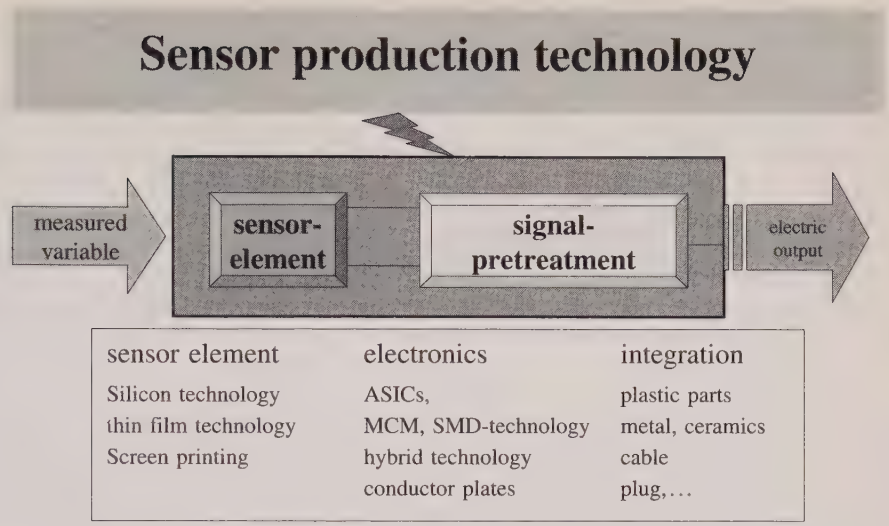


Figure 3.2.7. Sensor production technologies need a reliable and cost-effective integration of technologies for the sensor element, the electronic data (pre) processing assembly.

electronics toward miniaturization and increased complexity has also affected sensors.

The first silicon sensors were introduced in the fifties and sixties. They developed into semiconductor microsensors and in the past decade even microsystems.

Furthermore, other solid-state sensors have been developed and introduced, ranging from thin film and screen-printed sensors to ceramic sensors and – latterly polymeric sensors.

Progress could be made in the development of sensors where it was possible to sensors and (micro-)electronic integrated circuits ("chips") in a common encapsulation with compatible technology. So today we have various types of sensors on the market, including microelectronics on printed circuit boards, SMD-like devices or electronics on ceramic hybrid circuits as well as flexible polymeric boards. In some cases fully integrated intelligent silicon sensors are used where some data processing and a sensor are combined on one chip – or sometimes two chips. The decision which of these technologies is suitable depends on technical requirements for sensors and data processing, size and weight of the end product and not least on cost in relation to the numbers of the item produced.

The latest development are micromechanical sensors. Their development began with the large-scale introduction of silicon micromachined pressure sensors to the automotive industry in the nineties, which entailed a massive price reduction. Then acceleration sensors for airbag firing, yaw rate sensors and more were introduced. Many devices are still being discovered. The next step is product evolution, with introduction times between a few years and over a decade, as shown in Table 3.2.2. Once customers in the industry have accepted a product,

Table 3.2.2. Product evolution time table of selected MEMS sensors [7].

Product	Discovery	Product evolution	Cost reduction	Full commercialization
Pressure sensors	1954–1960	1960–1975	1975–1990	After 1990
Acceleration sensors	1974–1985	1985–1990	1990–1998	After 1998
Gas sensors	1986–1994	1994–1998	1998–2005	2005?
Chemical sensors	1980–1994	1994–1999	1999–2004	2004?
Yaw rate sensors	1982–1990	1990–1996	1996–2002	2002

investment in large-scale production can go ahead. It helps to find more applications for the product. The time scale for the product evolution process varies from about five years up to more than two decades, depending on acceptance and success as well as entrepreneurial initiative. Many of these reliable mass fabricated low cost devices were originally developed for the automotive industry and are now used in many other fields, e.g. the household industry.

We expect these micromachined sensors to become more and more important in the household industry, in many domestic applications of silicon pressure sensors, acceleration sensors, tilt sensors, infrared detectors and thermopiles, flow meters, as well as gas sensors and liquid constituent sensors.

Microsystems are also expected to be introduced in the near future, including for example artificial noses, fingerprint sensing systems, bar code readers, rf-tagging systems, microfluidic pumps and dosing systems, gas flow control systems, new flexible and low cost displays or electronic paper.

3.2.5 Conclusion

Several factors contribute to the total production costs of household appliances:

- Mechanical body construction between one quarter and half of the costs
- Electronics and power supply about one third
- Actuators about 20%
- Sensors 15% to more than 20%

As the domestic appliance industry caters for a mass market – the production volume of a single company may well be around a million units – interesting opportunities are opening up for sensor and microsystem producer, particularly, because silicon microsystems show their true potential especially when mass-pro-

duced. But there is also room for many other technologies in this interesting market, provided technology and cost are viable.

Many sensor–actuator control systems are already in use today, and many more will be introduced over the coming years, so that we can expect an increase in sensors, actuators and control systems, many of which will be microsystems, in the near future.

3.2.6 References

[1] Lahrmann, A., *Smart domestic appliances through innovations, 6th International Conference on Microsystems 98, Potsdam,* (Reichl, H., Obermeier, E., eds.), Berlin: VDE-Verlag, 1998.

[2] Appliance Magazine, *Portrait of the European Appliance Industry*, 2001.

[3] Stamminger, R., *Keynote for the components day at the Hometech, exhibition 28th February* 2002.

[4] *Zahlenspiegel des deutschen Elektro-Hausgerätemarktes,* edited from ZVEI, Zentralverband Elektrotechnik und Elektronikindustrie, Frankfurt: Hausgeräte-Fachverband, 2002.

[5] *Data collected during the meetings of the European NEXUS User-Supplier Club 7 Household Appliances* (see e.g. in *www.nexus-mems.com*).

[6] Tschulena, G., *mst news 2 (2001).* (Tschulena, G., ed.), Oxford: Commercialization of Microsystems, 2001.

[7] Grace, R., *Commercialisation Issues for Microsystems*, Proc. Microsystem Technologies, Düsseldorf, March 2001.

Index

A/D converters 60–80, 96 ff
absorption coefficients, optical 8
acceleration sensors 221, 240
acoustic mirrors 204
active alignment 146
active matrix arrays 5 ff, 27 ff
active matrix flat-panel imagers (AMFPIs) 4 ff
active pixel sensors (APS) 3 ff, 29
actuation
– thermomechanical 101
– cantilevers 102
– markets 231
– modular microoptical systems 145 ff, 157 ff
adhesives 167
air gaps, RF filters 183
alignments 146 ff, 151 f, 164
Allen variance 113
alloying 16
aluminum 12
aluminum nitride 188
aluminum/polysilicon thermopiles 77
amorphous selenium (a-Se) 6 f, 15 ff
amorphous silicon (a-Si) 3–50
amperometric methods 123
amplifiers
– active pixels 30 f
– calorimetric sensors 80
– chopper 51
– low-noise/ differential difference 51
analyte–polymer interactions 54 ff
angle sensors 224
angular rate sensors 223
annealing 9
antenna duplexer 177, 181
antialiasing filter 80, 96 ff
applications 141–210
arrays, nanoelectrodes 132
arsenic 16
a-Si:H array 39 ff
a-Si:H technology 4 ff, 8 ff, 19 ff
a-SiC:H 23
a-SiN_2:H 10
a-SiO_x (ON) 10
assembly technology 146
attenuation
– RF filters 177 ff, 181
– SCF 199
automotive sensor market 213–230
avalanche multiplication 7

back-channel conduction 28
balanced signals 181
baluns
– BAW filters 196
– RF filters 181 ff
bandpass filter 82, 88
bandwidth
– calorimetric sensors 81 ff, 93 f
– gas detection 73
barcode reading 166
BAW filters 194 ff
beam fiber coupler, expanded 149
beam size, cantilevers 102
biases
– active pixels 30
– X-ray imaging 19 ff
bioavailability 123
blocking contacts 15
blood flow sensor 155
Bode plots
– cantilevers 105
– DDA 110
boiling points 53

bonders
– modular microoptical systems 167
– thermal compression 158
bulk acoustic wave (BAW) 175–210
business models 216
Butterworth–van Dyke model 202
bypass switches 183

cadmium complexation 127
calibration, on-chip 99
calorimeters 51, 75 ff
calorimetric transducers 77
cantilevers
– gas detection 51
– modular microoptical systems 158, 162
– resonant beam sensor 100 ff
capacitance 41 ff, 85
capacitive crosstalk 106
capacitive currents 132
capacitors 57 ff
– RF filters 178
– tunable 183
carbon 9, 127
ceramic RF filters 178
channel resistance 67
charge-coupled devices (CCDs) 5 f
charge gain, active pixels 31
charge injection compensation 67 ff
charge transfer 132
chassis systems 228
chemical sensitivity 78
chemical sensors 51, 240
chlorine 16
chopper amplifier 165
– calorimetric sensors 80
– gas detection 51
– low-noise 81
chopping-clock generator 91
chopping frequencies 88
chromium 12, 32
circuit operation, active pixels 29
circuitry, DDA 111
clamping structures 165
climate control 231
clock cycles 62 ff
clusters
– ionomer-coated electrodes 127
– X-ray imaging 13
coated electrodes 121–141
coating 24, 36 ff
collimation optics 160
columnar growth 188
comb filters 97
comfort systems 219, 225
commercialization, BAW filters 206 ff
common-mode feedback (CMFB)
– amplifiers 108
– gas detection 73
common-mode rejection ratio (CMRR) 108
compact model, BAW filters 201
complementary metal oxide semiconductors (CMOS)
– gas detection 51–120
– processes 57
– switches 69
– X-ray imaging 5 ff
concentration, gas detection 56
conduction band 8 f, 20 f
construction principle, BAW filters 195
conversion efficiency, phosphor films 37
coplanar gate structures 27
copper complexation 127
correlated double sampling (CDS)
– calorimetric sensors 80 f
– capacity sensor 60
cost reduction 219
counters 74, 104
coupled acoustic electrical model, BAW filters 201
coupled resonator filters (CRF) 200
couplers, fiber 146 ff
coupling, effective 190
coupling coefficient, piezoelectric 189
cracking 32
cross connect 159 ff, 162 ff
crosstalk, cantilevers 106

dangling bonds 8 ff
dark currents 19, 25
decimation filters
– calorimetric sensors 97
– gas detection 51, 74 f
deep defects 8 ff
deep reactive ion etching (DRIE) 159
defect states 8 f, 20 ff
deflection, cantilevers 102
deformable mirrors 158
delay line, amplifiers 112
delayed switching 68
demodulation
– capacity sensor 60
– choppers 83
depletion 18
deposition methods 9
detection
– gas 51–120
– piezoresistive 101 ff
– X-ray imaging 6 ff, 15 ff
detection limits 125 ff, 132, 226
detuning 195
diagnostic methods, BAW resonators 193
dielectric constants 57
– BAW resonators 189
– imaging arrays 43
dielectric materials 10
differential difference amplifier (DDA) 51, 107
differential pulse voltammetry (DPV) 125, 128
diffraction 153
diffusion 21
diffusion coefficents 124
digital signal processors (DSP) 97
digital X-ray imaging, large-area 3–50
diode switches 183
dipole–dipole interactions 55
dispersion
– BAW resonators 193
– gas detection 55
distance sensors 154
donor model 9
doping 9 f, 23
Doppler method 155
double differential methods 132
drain current, active pixels 31
drivetrain 219, 227
dryers 234
dual-gate structures 27
dummy switch compensation 68
duplexer filters 183
duplexer systems 169

Eastman AQ55 124 ff, 131 ff
effective coupling, BAW resonators 190
electrical acoustic domains 179
electrical characteristics 3 ff
electrostatic spray deposition 130
electroactive substances 122
electromagnetically actuated laser scanner 166
electron diffraction, ITO 15
electron–hole pairs 6, 15 ff, 28 ff
electronic noise 70
electrooptical components, modular 151, 157 ff
electroplating 164
electrostatic discharge (ESD) 179
electrostatically actuated cross connects, optical 159
engine combustion sensors 227
enthalpy change 76
environmental monitoring 121–141
equivalent circuit 58 ff
error amplifier 69
etching 10, 152, 159
ethylene glycol-based polymer 36
European markets 233
excitation signals 101
exhaust gas sensors 224
expanded beam fiber coupler 149

Faradaic current signals 123, 133
feedback
– cantilevers 107

– interdigitated capacitors 64
– resonant beam sensor 101
Fermi level 20
ferrules 148
fiber couplers 146
fiber switches, hybrid 161
field-driven MEMS resonators 184
figure-of-merit, BAW resonators 190
film bulk acoustic wave resonator (FBAR) 185
film thickness, coated electrodes 130
filters
– BAW 194 ff
– calorimetric sensors 82, 88, 96 ff
– thin films 175–210
fingerprint sensors 59
finite-element method (FEM) 58, 200
finite-impulse response (FIR) 97 ff
fixation methods 146
flexural mode devices 185
flicker noise
– calorimetric sensors 86
– gas detection 74
flip-chip assembly 143, 158 ff, 170 f
floating potential, large-area arrays 42
fluoroscopy 7
frequencies
– BAW resonators 190
– chopping 88
– resonant beam sensor 101
– RF filters 177
– temperature coefficient (TCF) 191
Fresnel reflection 153
fringing factor 43
front-channel conduction 28
fully-differential modulators 68
functionality 220

gain
– amplifiers(DDA) 109 f
– calorimetric sensors 81, 93
gain-bandwith product (GBW) 73
gain resolution 37
gaps, X-ray imaging 8 f, 35
gas detection, CMOS single chip 51–120
gas sensors 240
gate bias 5 ff
gate metals 12 ff
gate structures 27
Gd_2O_2S:Tb phosphor film 35 ff
geometric overlap capacitance 43
Gibbs energy 55 ff
glass substrate 24
glue, fiber couplers 147, 151, 161
gold connections 152
gold layers 154
grains, BAW resonators 193
grating structures, RF filters 180
grippers 167
grooves 147 ff, 161

handling conditions 167
harmonic modes 188
heating 231
heating-power deflection 106
heterodyne receiver 151
high-pass filter 113
high-Q inductors 183
hillock generation 12
household appliances 231–241
hybrid-fiber switch 161
hybrid microoptical systems 143–174
hydrogen bonds 55
hydrogen incorporation 8

illumination
– active pixels 30
– X-ray imaging 26
imaging, large-area digital X-ray 3–50
impedance level
– BAW filters 194, 201
– BAW resonators 189
– RF filters 181 ff, 184
impurities 9
indicator electrodes, ionomer-coated 129

inductors 178, 183
infrared light 164
injection-molded fiber coupler 149
integrated components, passive 182
integrated sensors 51
integrated systems, modular microoptical 146
integration capacitors 62
integration mode, active pixels 30
interactions, analyte–polymer 54 ff
interdigitated capacitors 57, 64
interfaces
– capacity sensor 60 ff
– gas detection 51
intermediate frequency filter (IF) 177, 181
inversion, calorimetric sensors 84 f, 87 f
inverted staggered gate structures 27
ion-exchange voltammetry (IEV) 121–141
ionomer-coated electrode ensembles 121–141
islands, n-well 65, 77
ISM bands (industrial, scientific, medical) 178
ITO/a-Si:H diodes 16

Kapton substrate 44

ladder BAW filters 196
Langmuir–Blodgett films 130 ff
large-area digital X ray imaging 3–50
laser diode 155
laser scanner 166
lattice BAW filters 198
laundry 231
layers
– calorimetric sensors 77
– gas detection 53
– ionomer-coated electrodes 130 ff
– modular microoptical systems 154
– resonant beam sensor 100
– RF filters 186 ff
– X-ray imaging 17, 32
lead complexation 127
lead iodide (PbI_2) 15
lead zirconium titanate (PZT) 185, 188 f
leakage currents 3, 19, 27 ff, 33
lens arrays 150
LIGA process 143–174
ligands 122
light generation/detection 146
linear motion, cross connections 159
linearity, active pixels 30
liquid-media condition control 227
lithium niobate/tantalate 180
London forces 55
loop filter 61
low-noise amplifiers 51
– chopper 80 f
– differential difference (DDA) 107 ff
low-noise OTA 72
low-temperature processes 3 ff, 5 ff

magnetron sputtering 8, 189
mammography 7
markets 211–241
Mason model 188, 201
mass flow sensors 223
material coupling 193
material losses 189
material selection
– piezolayers 188
– X-ray imaging 7 ff, 12 ff
medical imaging arrays 7
membranes
– BAW resonators 186
– calorimetric sensors 77 f
– ionomer-coated electrodes 121
mercury drop electrode 123
mercury speciation 128
metal indicator electrodes 129
metal-oxide semiconductor (MOS) 68
metastability 9
methylmercury speciation 128

microchoppers 165
microelectromechanical systems (MEMS) 145 ff, 159, 175–215
microoptoelectromechanical system (MOEMS) 145 f
microoptical systems, modular 143–174
microsensor interfaces 60
microsystems markets 213, 231
microwave filters 177
microwave ovens 234
Miller capacitator 109
miniaturization 219
mirrors
– BAW 187, 204
– modular microoptical systems 157 f
Mo/a-Si:H diodes 7, 16 f
mobile phones 177, 181 ff
mobility 15 f
modeling, BAW filters 200
modular microoptical systems 143–174
modulation transfer function (MTF) 37
modulators
– gas detection 68
– modular microoptical systems 164
– $\Sigma\Delta$ 60
molded fiber coupler 149
molecular engineering 130 ff
molybdenum 12 f, 24, 32
monolilthic systems, modular microoptical 146
monolithic microwave ICs (MMICs) 226
monolithic sensors 51
motor control 227
multidielectric media, large-area arrays 42
multilayer structures 32
multiple square-wave voltammetry (MSWV) 125, 132
multipliers 97
multiply charged ions 121
multipole expansion 42
multisensor chips 72
n-type switches (NMOS) 69, 86
n-well 57, 65, 77
N×N switches 162 f
Nafion 124 ff, 131 ff
nanoelectrode ensembles 121–141
negative electrostatic spring 184
network analyzer 193
nickel 12
night-vision systems 227
nitrogen 9
nitrosamines 126
noise
– calorimetric sensors 81 f, 86 f, 95 f
– reference capacitors 64
– resonant beam sensors 101
– X-ray imaging 19
noise power transfer function (NTF) 70
nonlinearity, interdigitated capacitors 65
Nyquist A/D converters 71, 96

ohmic conduction 28
ohmic switches 183
on-chip calibration 99 f
on-chip circuitry, DDA 111
open-circuit potential 130
operating principles, X-ray imaging 3 ff
operational amplifier 71
operational transconductance amplifier (OTA) 62, 68, 70
optical absorption coefficients 8
optical benches 151, 157 ff
optical characteristics, X-ray imaging 3 ff
optical lithography 152
optical systems, modular 145
optimum thicknesss 19, 32
optoelectromeachanical systems 151 ff, 157 ff
optomechanical systems 146
organics
– ionomer-coated electrodes 122
– X-ray imaging 36

original equipment manufacturers (OEM) 216
oscillators
– calorimetric sensors 88
– infrared light 164
– resonant beam sensor 101, 107
ovens 234
oversampling ratio 60
oxygen 9, 21

p-i-n diodes 3 ff, 15 ff, 23 ff
p-type switches (PMOS) 69, 86
packaging, BAW filters 206
parallel plate approximation 43
parasitic capacitors
– gas detection 62, 71
– X-ray imaging 3 ff
parasitic substrate effect 191
parathion 126
partition coefficient
– calorimetric sensors 78
– capacity sensor 59
– gas detection 55
passivation
– gas detection 57
– X-ray imaging 8
passive assembly 146
passive optical chip 155
passive pixel sensor (PPS) 27
patterning
– modular microoptical systems 158
– X-ray imaging 32
Permalloy 164 ff
phase margin, DDA 110
phase sensors 223
phenol 126
phosphor-coating solution 36
phosphor integration 35 ff
phosphor screen 5 ff
photodiodes 151, 155
photoelectric effect 15
physical sensitivity, calorimetric sensors 78
physical vapor deposition (PVD) 7
physisorption 55 ff
piezoactuators 160
piezoelectric detection 103
piezoelectric substrates 180, 186
piezoresistive detection 101 ff
pin insertion 148
pixels 158
– integration 27 ff, 32 ff
– structures 5 ff, 27 ff
plasma-enhanced chemical vapor deposition (PECVD) 7
plastic substrates, flexible 44
plate capacitors, polymer-filled 57
platforms 229
pollutants 123
poly(dimethylsiloxane) (PDMS) 55
poly(etherurethane) (PEUT)
– cantilevers 105
– gas detection 55, 58
poly[methyl(cyanopropyl)siloxane] (PMCPS) 55
poly(methyl methacrylate) (PMMA) 148, 154
poly(methyloctylsiloxane) (PMOS) 55
poly(vinyl alcohol) (PVA) 36
polycrystalline layers 188
polyester sulfonate 125
polyimide waveguide 156
polymer–analyte interactions 54 ff
polymer layers
– calorimetric sensors 77
– ionomer-coated electrodes 130 ff
– resonant beam sensor 100
polymers 105
– fiber couplers 148
– gas detection 51 ff
polysilicon 65
polysiloxane derivatives 55
position sensitive detector (PSD) 154
position sensors 224
potential fluctuation model 9
power consumption
– choppers 94
– DDA 110
power handling, BAW resonators 192

preamplifier, low-noise 83
pressure sensors 222, 240
processing
– CMOS 57
– X-ray imaging 3 ff
production figures 231
pyramidal grooves 154

quality, markets 220
quality factor
– BAW resonators 187, 190, 194
– resonant beam sensor 101, 105
– RF filters 177
quantization noise 70
quartz crystals
– BAW filters 202
– RF filters 180, 186

radar sensors 225
radiography 7
Rayleigh waves 179
read-out circuitry
– calorimetric sensors 80
– capacity sensor 59
read-out mode
– imaging arrays 40
– active pixels 30
receive bands, RF filters 177
receiver, heterodyne 151
recombination 8
redox states 121 ff, 129 ff
reference capacitor, tunable 63
reflection grating 154
reflectivity 187
refractory metals 12 ff
refrigerators 234
relaxation oscillators 60
reliability, BAW filters 207
Remez exchange algorithm 98
reset mode, active pixels 30
resistors 65
resonance frequencies 102
resonant-beam gas sensors 100
resonant-beam oscillators 107
resonators
– BAW 184
– RF filters 179 ff, 183 ff
reverse currents 28
RF filters, MEMS 175–210
rhombus-like grooves, silicon 147
ribbons, fiber couplers 150
Richardson constant 20
robustness, BAW resonators 188
rotating mirrors 162
rotational speed sensors 223

safety systems 219, 225
scaling factor, choppers 84
scanners, barcode reading 166
scanning electron microscope (SEM)
– fiber couplers 148
– gas detection 57
– modular microoptical systems 158
Schottky barriers 20 ff
Schottky diodes 3–18
Schubweg 15 f
screen film systems 5 f
$\Sigma\Delta$ A/D converters 96 ff
$\Sigma\Delta$ modulator 51, 60 ff
Seebeck coefficient 78
segmented deformable mirrors 158
selenium technology 6 ff, 15 f
semiconductors manufacturers 217
sensing principle
– calorimetric 75 f
– capacitive 57
sensitive layers, polymers 53 ff
sensitivity
– resonant beam sensor 100
– calorimetric sensors 78
sensor specialists, markets 217
shear stress, fiber couplers 147
short-range radar systems 226
shunt resonators, BAW 196
signal-dependent charge injection 68
signal preamplification 66

signal-to noise-ratio (SNR) 70 ff, 79
– calorimetric sensors 97
– ionomer-coated electrodes 134
signal transfer function, choppers 82
signals, cantilevers 106
silicon
– cantilevers 100 ff
– fiber couplers 147
– n-well 57
– technology 3–50
simulation, DDA 110
sinc^l-filters 97 ff
single-bit $\Sigma\Delta$ modulator 60
single-chip gas detection 51–120
single-crystal silicon 147
small-signal amplitudes 106
Smith chart 204
sol-gel processes 190
solidly mounted BAW resonators 187
sound velocity 189
source drain 28
spacing, substrate/n-well island 79
spectrometers 153
spurious modes, BAW resonators 192 ff
sputtering
– BAW resonators 189
– X-ray imaging 8, 12, 17
stability
– cantilevers 113
– gas detection 55
– X-ray imaging 4, 19 f
stacked BAW crystal filter (SCF) 199
stacked structures
– BAW filters 196
– pixels 32
– X-ray imaging 12
Staebler–Wronski effect 9
staggered gate structures 27
state-of-the-art BAW filters 203
steerable mirrors 157 f
stopband attenuation 181
stress-sensitive MOS transistors 103
strong inversion, choppers 82
substrate effects 191
sulfonated fatty acids 36
sulfonated polyesters 124 ff
surface active agents 122
swelling 59
switched capacitor amplifier 62
switching
– cross-coupled 82
– gas detection 68
– modular microoptical systems 159
– RF MEMS 182 ff
– rotating mirrors 162 f
– X-ray imaging 5 ff
system in chip (SIC) 207
system on chip (SOC) 207
system suppliers 216

telecommunication 143–174
temperature coefficient (TC)
– BAW filters 204
– BAW resonators 189
– gas detection 64
temperature increase, calorimetric sensors 76
template synthesis 121
thermal compression bonding 158
thermal crosstalk 106
thermal noise
– calorimetric sensors 86
– cantilevers 113
– gas detection 74
thermally-insulated n-well islands 77
thermocouples 79
thermodynamic equilibrium 76
thermoionic emission 20
thermomechanical actuation 100, 104
thermopiles, aluminum/polysilicon 77
thin-film bulk acoustic wave technology 175–210
thin-film techology 7, 10
thin-film transistor (TFT) 4 f, 10 ff, 27 ff, 40 ff
threshold voltage, X-ray imaging 4 ff
tin-doped indium oxide (ITO) 11 ff
topologies, capacity sensor 59
Tosflex 124 ff, 131 ff

toxicity 121 ff
trace concentrations 121
transconductance 84
transducers 55
transductors 90
transfer function
– DDA 110
– signals 82
transient slope, switching 68
transistor size 110
transmission electron microscope (TEM) 11
transmission line 201
transmit bands 177
transmitter module 171
trapping 8, 15, 21 f
trends, markets 213, 220
triangulation principle 154, 161
trimethylborane (TMB) 24
tunable reference capacitor 63
tuning fork resonator 180
tunneling 8, 20

ultra-high vacuum (UHV) 9
ultratrace concentrations 121
UV curing adhesives 148, 151, 161

V-grooves 147, 155
vacuum tube 6
valence band 8
vertical structures, TFTs 44
vibration amplitudes 194
video sensor markets 225
viologen 126
voids 9
volatile organic compounds (VOCs) 53
voltammetry 121–141
volume fraction 59

wafer-level package 207
washing machines 234
wavelength filter 151
weak inversion 87
wet etching 152
Wheatstone bridge 102 f, 106 f, 114
wireless local area networks (WLAN) 177
working principle, ladder filter 197

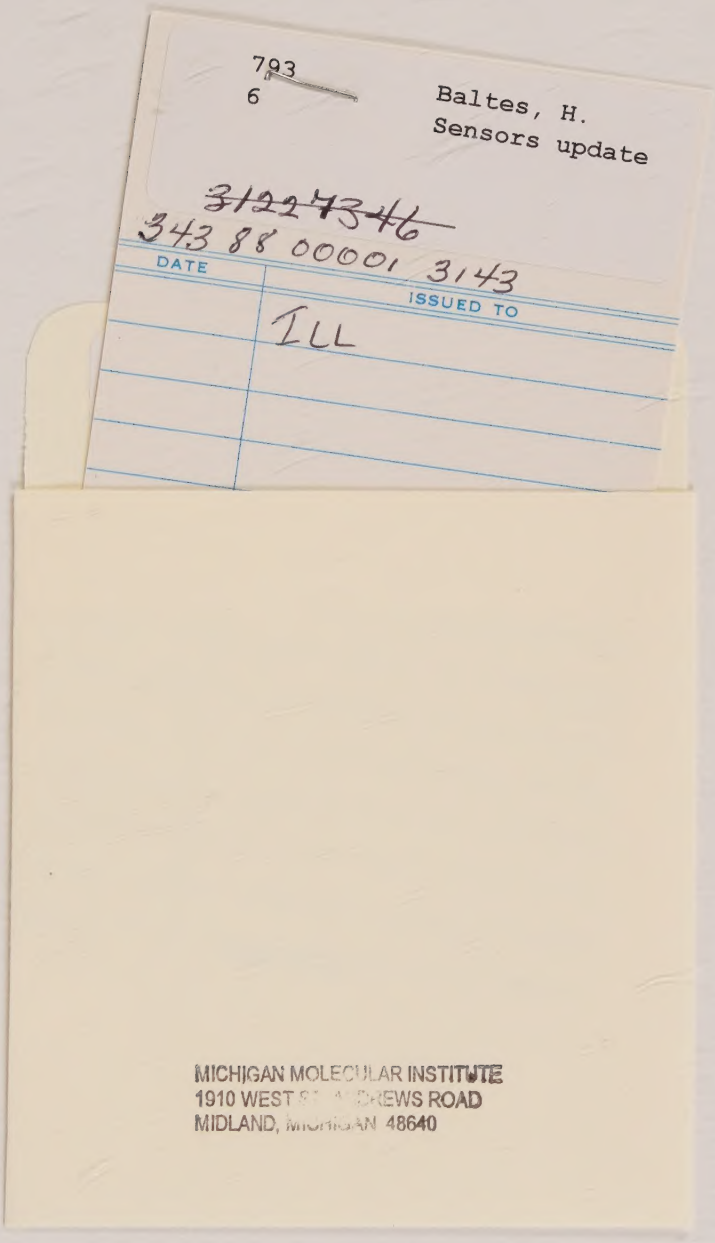
X-ray diffraction (XRD) 193
X-ray imaging, large-area digital 3–50

yaw rate sensors 240

zinc oxide 188

MICHIGAN MOLECULAR INSTITUTE
1910 WEST ST. ANDREWS ROAD
MIDLAND, MICHIGAN 48640

MICHIGAN MOLECULAR INSTITUTE
1910 WEST ST. ANDREWS ROAD
MIDLAND, MICHIGAN 48640


793
6
Baltes, H.
Sensors update
3122 7346
343 88 00001 3143
DATE
ISSUED TO
ILL
MICHIGAN MOLECULAR INSTITUTE
1910 WEST ST. ANDREWS ROAD
MIDLAND, MICHIGAN 48640

MICHIGAN MOLECULAR INSTITUTE
1910 WEST ST. ANDREWS ROAD
MIDLAND, MICHIGAN 48640

WILEY-VCH

MICHIGAN MOLECULAR INSTITUTE
3 4388 00001 3143

Sensors

Update

Edited by
H. Baltes, G. K. Fedder, J. G. Korvink

Volume 12

Sensors Update includes information at the cutting edge of sensor research and related topics. An overview of highlights in this exciting area is presented that covers current developments in materials, design, production and application of sensors and other micro- and nanosystems.

Volume 12 gives you an insight into current research and application fields of environmental sensors, integrated gas detectors, and X-ray imaging. Recent advances in the broad and important application areas of micro-optical and RF-microelectromechanical systems are reported and the sensor markets section features overviews for the automotive and household appliance sectors.

The unique approach makes *Sensors Update* a valuable reference work for scientists and engineers, developers, distributors and users of sensors in industry and at university.

www. sensorsupdate.com
www. wiley-vch.de

ISSN 1432–2404

ISBN 3-527-30602-1

9 783527 306022